December 2002

Representing Part-Whole Relationships in Conceptual Modeling: An Empirical Evaluation

Graeme Shanks
The University of Melbourne

Elizabeth Tansley
Central Queensland University

Jasmina Nuredini
The University of Melbourne

Daniel Tobin
The University of Melbourne

Ron Weber
The University of Queensland

REPRESENTING PART-WHOLE RELATIONSHIPS IN CONCEPTUAL MODELING: AN EMPIRICAL EVALUATION

Graeme Shanks
The University of Melbourne
Parkville, Victoria AUSTRALIA
g.shanks@dis.unimelb.edu.au

Elizabeth Tansley
Central Queensland University
Rockhampton, Queensland AUSTRALIA
e.tansley@cqu.edu.au

Jasmina Nuredini
The University of Melbourne
Parkville, Victoria AUSTRALIA
jasminan@staff.dis.unimelb.edu.au

Daniel Tobin
The University of Melbourne
Parkville, Victoria AUSTRALIA
danielt@staff.dis.unimelb.edu.au

Ron Weber
The University of Queensland
Brisbane, Queensland AUSTRALIA
weber@commerce.uq.edu.au

Abstract

The part-of or part-whole construct is a fundamental element of many conceptual modeling grammars that is used to associate one thing (a component) with another thing (a composite). Substantive theoretical issues surrounding the part-whole construct remain to be resolved, however. For instance, contrary to widespread claims, the relationship between components and composites is not always transitive. Moreover, how the part-whole construct should be represented in a conceptual schema diagram remains a contentious issue. Some analysts argue composites should be represented as a relationship or association. Others argue they should be represented as an entity. In this paper we use an ontological theory to support our arguments that composites should be represented as entities and not relationships or associations. We also describe an experiment that we undertook to test whether representing composites as relationships or entities enables users to understand a domain better. Our results support our arguments that using entities to represent composites enables users to better understand a domain.

Keywords: Conceptual modeling, information systems development, ontology, part-of relationship, aggregation, composition, meronymic relations, mereology, mereotopology

1. INTRODUCTION

The notion that one thing may be a *part of* or a *component of* another thing (e.g., a wheel is part of a car) seems fundamental to the way humans conceive some types of phenomena in the perceptual worlds they create. For this reason, part-whole (meronymic) relations have been a focus of psychologists concerned with human cognition (e.g., Winston et al. 1987) and philosophers concerned with ontology (the nature of the world) (e.g., Bunge 1977). Indeed, within the field of philosophy, part-whole relations are the primary focus of the subfields of mereology or mereotopology (e.g., Simons 1987, Smith 1996, Varzi 1996).

Part-whole relations also have long been a concern of information systems researchers and practitioners concerned with finding better ways to model the world as a basis for building better information systems. For instance, they feature in early work on "database abstractions" (Smith and Smith 1977) and extensions to Chen's (1976) entity-relationship model (Teorey et al. 1986). Subsequent to this early work, they have remained a focus of researchers concerned with conceptual modeling (e.g., Storey 1991). More recently, part-whole relations feature in object-oriented conceptual modeling approaches (e.g., Opdahl et al. 2001)—in particular, the *de facto* standard for object-oriented conceptual modeling, the Unified Modeling Language (UML) (Rumbaugh et al. 1999).

In the context of conceptual modeling work in the information systems discipline, part-whole relations remain problematical for two reasons. First, substantive theoretical issues surrounding part-whole relations remain unresolved. For example, Rumbaugh et al. (1999, p. 146) state: "The aggregation (part-whole) relationship is transitive and antisymmetric across all aggregation links, even across those from different aggregation associations." Winston et al. (1987, pp. 431-432) illustrate the difficulties with such claims, however, and a deeper understanding of part-whole relations is needed as a basis for conceptual modeling languages and methods (Wand and Weber 2002).

Second, alternative ways of representing part-whole relations in conceptual models have been proposed. In particular, composite things are sometimes represented explicitly as entities (e.g., Kilov and Ross 1994, pp. 96-97) and sometimes represented implicitly via relationships between the components of the composite (e.g., Chen 1976, p. 31). The merits of these alternative representations have been evaluated theoretically (Wand et al. 1999). To the best of our knowledge, however, no rigorous empirical evaluation of alternative representations of part-whole relations has been undertaken.

In this paper, therefore, we describe research we undertook to empirically evaluate alternative conceptual-modeling representations of part-whole relations. Our motivation was fourfold. First, it is well known that the cost of fixing errors grows exponentially the later they are discovered in the system development process (e.g., Boehm 1981). Because conceptual modeling work is undertaken early in the system development process, improvements in conceptual modeling practice potentially will lead to high payoffs (Moody and Shanks 1998).

Second, practically it is important to determine which type of representation of certain real-world phenomena enables humans to understand the phenomena better. When conceptual models are first prepared (e.g., by systems analysts), the users of an information system are asked to evaluate the models to determine how accurately and completely they represent the users' perceptual worlds. If users cannot understand the conceptual model clearly, their ability to validate the model will be impaired. Moreover, subsequently users may employ conceptual models to try to understand the functionality provided by an information system. Again, if users cannot understand the conceptual model clearly, their ability to comprehend and use the information system will be impaired.

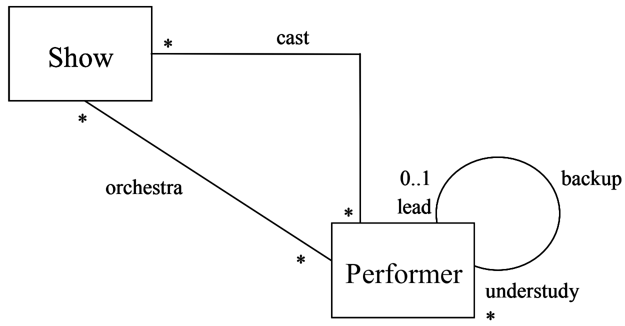
Third, we sought to test prior theoretical work undertaken to predict how well different types of representations facilitate or inhibit human understanding of real-world phenomena. If we can make accurate predictions about what types of conceptual modeling practices are likely to work well, we avoid the high costs associated with learning about the strengths and weaknesses of different practices through experience.

Fourth, we sought to contribute to improved conceptual modeling practice. Part-whole relations employ some form of the relationship construct in conceptual models. The relationship construct is used either to represent the composite implicitly or to link components with their composite explicitly. Prior research has shown, however, that database designers often find the relationship construct difficult to use appropriately (e.g., Batra et al. 1990; Goldstein and Storey 1990; Prietula and March 1991). If we can develop improved conceptual modeling rules for part-whole phenomena, we will assist practitioners.

2 THEORY AND PROPOSITION

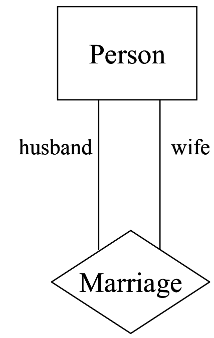
Figures 1 through 4 show some examples of how part-whole relations have been represented in some important conceptual modeling literature. Figures 1 and 2 show a composite represented *implicitly* via a relationship construct, and are drawn from widely used sources. Figures 3 and 4 show a composite represented *explicitly* via an entity construct, and are alternative representations that we propose. Which type of representation is "better"? Does it matter which is used?

Wand et al. (1999) contend that Figures 1 and 2 are the poorer representations, while Figures 3 and 4 are the better representations. Their arguments are founded on an ontological theory proposed by Bunge (1977). In brief, their analysis runs as follows:



(Source: Rumbaugh et al. 1999, p. 154)

Figure 1. Backup Composite Represented as a UML Association



(Source: Chen 1967, p. 31)

Figure 2. Marriage Composite Represented as an E-R Relationship

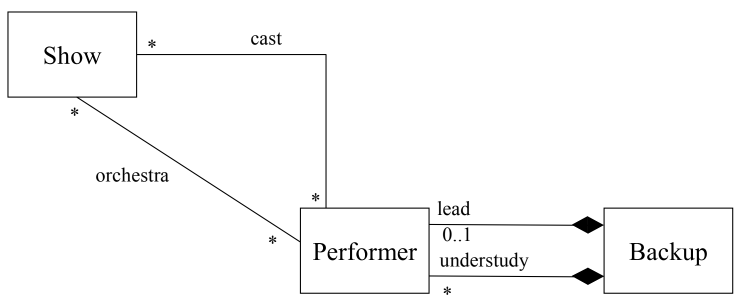


Figure 3. Backup Composite Represented as a UML Class

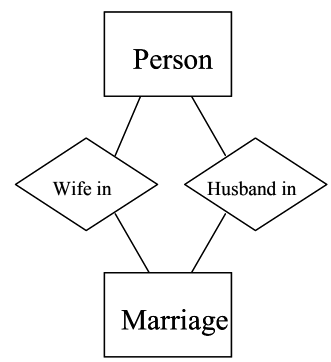


Figure 4. Marriage Composite Represented as an E-R Entity

1. "The world is made of things that possess properties" (p. 497). Things and properties are the two atomic constructs needed to describe the world.

2. Every thing in the world possesses one or more properties (there are no bare things) (p. 498).

3. Properties themselves cannot have properties. Moreover, properties cannot exist by themselves. They must attach to some thing (p. 498).

4. Two types of properties that exist in the world are *intrinsic properties*, which depend on one thing only, and *mutual properties*, which depend on two or more things (p. 498).

5. Two things interact (are coupled) when a history of one thing (manifested as a sequence of the thing's states) would be different if the other thing did not exist (p. 503)

6. The existence of a mutual property between two things can indicate that they interact with each other. Mutual properties that manifest interactions between two things are called *binding mutual properties* (p. 503).

7. "Two things may associate to form another thing." A thing is a *composite* if and only if it is formed from the combination of at least two other things. Otherwise, it is a *simple* thing (p. 504).

8. Every composite thing possesses emergent properties—properties not possessed by the components of the composite (p. 504).

In the context of Bunge's ontological theory, a composite can *not* be represented as a relationship. The reason is that relationships themselves represent mutual properties (4 above). Every composite, however, must possess at least one emergent property (8 above). If nothing else, a composite possesses the emergent binding mutual property of being related to its components. Its history depends on the histories of its components. If a composite is represented as a relationship, therefore, the relationship (mutual property) must itself have a property associated with the existence of at least one emergent property. Recall, however, that only things possess properties—properties themselves can *not* possess properties (3 above). Thus, composites must be things.

If the ontological principles are contravened and composites are represented as relationships, we argue that the resulting conceptual schema diagram is limited in several ways. First, users of the diagram must employ tacit knowledge to determine whether the relationship represents a composite thing or a mutual property of (relationship between) two or more things. For example, in Figure 1, the label "backup" could be interpreted as a mutual property or relationship (association) between "performer" things. It could also be interpreted as a composite thing called "backup" with components performer (in the lead role) and performers (in the understudy role). The composite thing "backup" would then have a separate existence with its own properties (for example, the size of each backup team and the date it was created). Unfortunately, like many other conceptual modeling grammars, UML permits construct overload to occur whereby the relationship construct sometimes represents an entity and sometimes represents a mutual property. Users of the diagram then have to employ tacit knowledge to determine what ontological construct is being represented by the grammatical construct (Wand and Weber 1993). Admittedly, the usefulness of distinguishing between things and properties remains a moot issue (Halpin 1995). Nonetheless, in the conceptual modeling domain, some research indicates that humans distinguish between things and properties as a way of managing complexity in real-world phenomena they are seeking to understand (e.g., Moody 2001; Weber 1996). Sustaining this distinction in conceptual schema diagrams, therefore, should help users to better understand the phenomena the diagrams are intended to represent.

Second, if *intrinsic* properties are attached to the relationship, are these properties intended as properties of the relationship or properties of the composite? For example, in Figure 1, assume that an attribute called "critic's rating" is attached to the "backup" relationship. Does the critic's rating apply to the role that *individual* performers play as a lead or understudy? Or is it an emergent property that applies to the composite backup team comprising both the lead and backup parts? In short, with the composite represented as a relationship, it is not clear whether the critic's rating applies to a component of the composite or the composite itself.

Third, in a similar vein, if *mutual* properties are attached to the relationship, are these properties intended as properties of the relationship or properties of the composite? For example, in Figure 1, assume that a relationship called "managed by" is attached to the "backup" relationship and to another entity class called "agent."[1] Does the "managed by" mutual property apply to the relationship that *individual* performers have with an agent? Or does it apply to the relationship that the composite backup team has with an agent? In short, does the agent manage the backup team or individual performers in the backup team?

Fourth, if composites are represented as relationships, subclasses of the composite cannot always be shown easily. For example, in Figure 2, assume for purposes of determining taxation liabilities and social security benefits that we wish to represent whether married couples have children and, if so, whether these children have certain properties. The first step is straightforward: we might attach an attribute to the "marriage" relationship showing "number of children" with the value zero if the couple have no children. However, we might then want to show whether some children have disabilities, some are dependents, some have a record of juvenile offences, and so on. Through judicious use of optional properties, we can represent these phenomena. As more optional properties are used, the "laws" and behavior that apply to couples and their children become increasingly difficult to specify and to comprehend. Such difficulties motivate the use of subclasses (Wand et al. 1999). If "marriage" is shown as a relationship, however, subclasses cannot be used.

Fifth, often the semantics that apply to the phenomena are unclear. For instance, in Figure 1, can a performer be a lead in one backup and an understudy in another backup? The UML *multiplicities* show that a performer may or may not be a lead in a backup. If they are a lead, however, they can be a lead in only one backup team. On the other hand, a performer can be an understudy in zero to multiple backups. Nonetheless, the multiplicities do not enable us to determine whether being a lead in one backup precludes a performer from being an understudy in another backup. Moreover, UML seems to provide no way to represent this constraint in the diagram (other than including a text comment). On the other hand, if "backup" were represented as an entity class, and performers were divided into "lead" and "understudy" subclasses, this constraint could be shown easily (Figure 3).

[1]In the entity-relationship model, relationships cannot be attached to relationships. To circumvent this problem, often *objectified relationships* are used. Nonetheless, some conceptual modelling grammars allow relationships to be attached to relationships (Elmasri and Navathe 2000, pp. 103-104).

In light of our arguments above, we contend that the choice of representation for composites and components is important in terms of users' ability to elicit the meaning of the phenomena described via the representation. Hence, the following proposition motivates the empirical work we undertook:

> *Proposition*: Conceptual schema diagrams that use an entity class construct to represent a composite will enable their users to better understand the semantics associated with the composite than conceptual schema diagrams that use a relationship class construct to represent the composite.

3 RESEARCH METHOD

An experimental setting was chosen for this research to control for extraneous factors that might confound any impacts of alternative representations of part-whole relations on how well users understand these relations.

3.1 Design and Measures

A two-group, post-test only experimental design was used with one active between-groups factor. This factor, "type of representation," had two levels. The first, the ontologically sound level, had both composites and components in part-whole relations represented as entity classes in a UML class diagram. An association construct was used to show the relationship between composites and components. The second, the ontologically unsound level, had components represented as entity classes in a UML class diagram. Composites were implied, however, via associations between components.

The dependent variable, performance, was evaluated using the participants' problem-solving performance. Relative to recall and comprehension tasks, problem-solving tasks seem to provide a better indicator of someone's "deep" understanding of a domain (see, e.g., Mayer 1989). Following Gemino (1999) and Bodart et al. (2001), therefore, we used problem-solving tasks as a means of testing how well conceptual models communicate the semantics of a domain to users. We measured problem-solving performance in three ways: (1) solution accuracy, (2) time taken to provide a solution, and (3) normalized accuracy (total solution accuracy score divided by total time taken). For each problem, solution accuracy was evaluated in terms of whether participants obtained a correct answer to the problem, used tacit knowledge to obtain an answer, provided a clear explanation of their rationale, and interpreted the model correctly. Normalized accuracy was used to take into account that problem solvers might trade off solution accuracy against the time taken to perform a task.

3.2 Materials

Four sets of materials were used in the experiment. The first was a summary of the UML symbols used in the diagrams provided to participants in the experiment.

he second set of materials comprised two UML class diagrams (diagrams that provide a static view of phenomena) of a project-planning domain. The first showed an ontologically sound conceptual model (Figure 5) in which both components and composites were represented explicitly as entity classes. The second showed an ontologically unsound conceptual model (Figure 6). In other words, composites were implied via links between component classes. The class diagram was sufficiently rich to make some problem-solving tasks difficult. For both diagrams, data dictionaries were also prepared to show the definition of attributes notated on the diagrams.

The third set of materials comprised 11 problem-solving questions. As much as possible, the questions were designed to force participants to use the UML class diagrams to obtain a correct answer rather than rely on tacit knowledge of the project-planning domain. Responses to questions were "possible," "not possible," or "not sure" with a brief explanation. Four example questions, including two that led to non-significant outcomes (questions 1 and 2) and two that led to significant outcomes (questions 3 and 4), are:

1. Project X is made up of 10 phases. Does the model allow more than one team leader to work on some of the phases?

2. A team leader has resigned. Does the model allow the team to continue to work on the project without him?

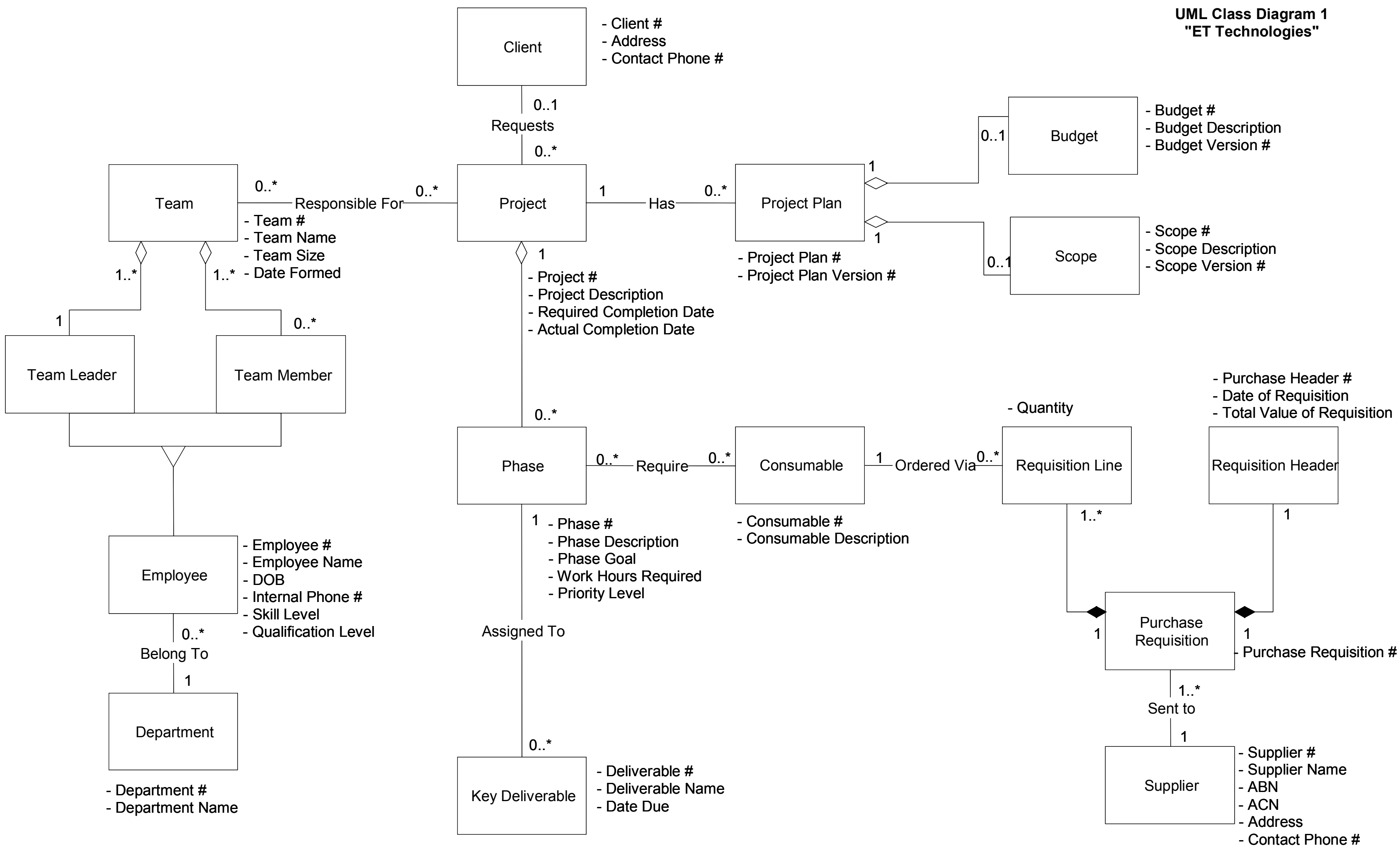


Figure 5. Ontologically Sound UML Class Model

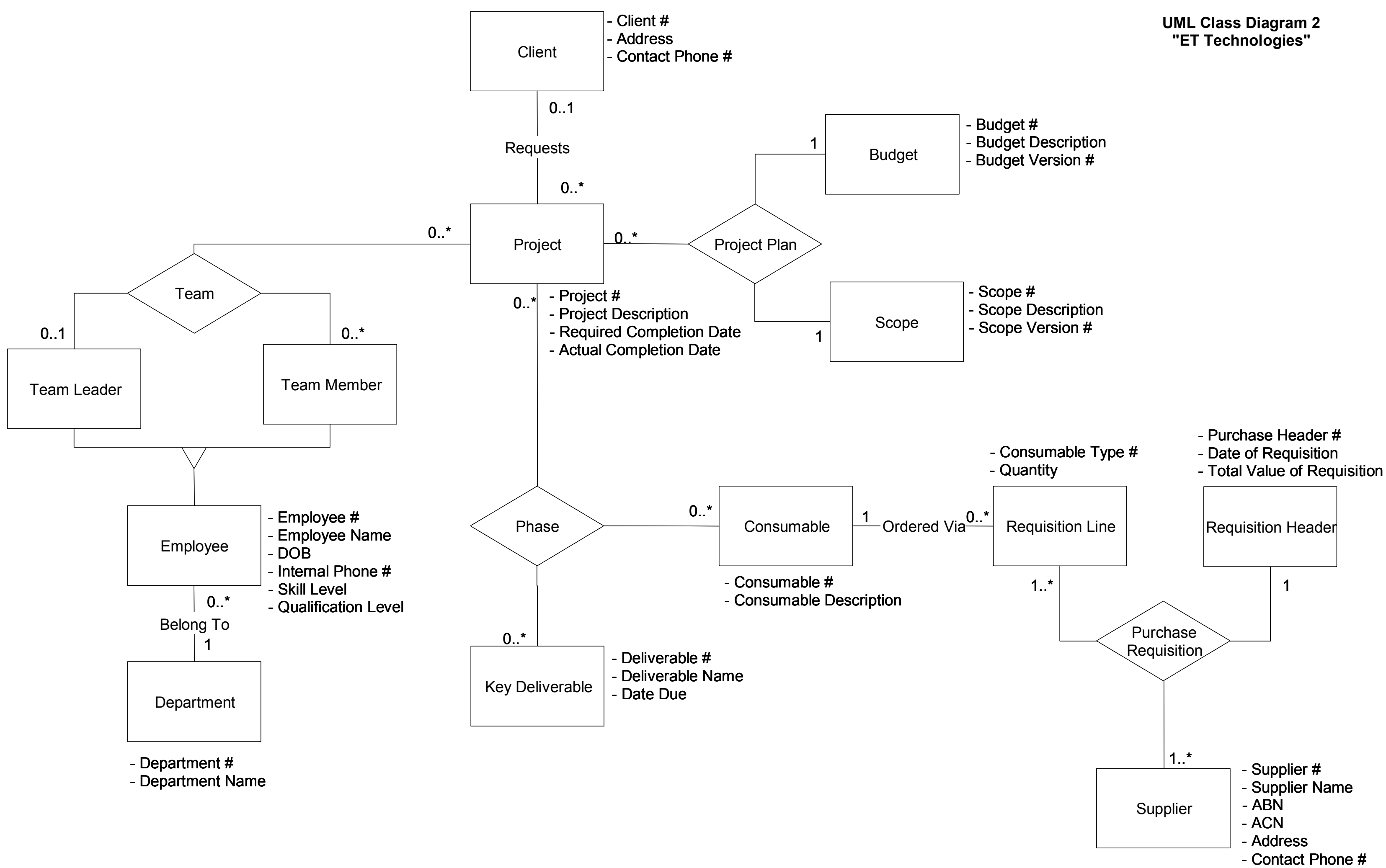


Figure 6. Ontologically Unsound UML Class Model

3. A client has requested that a project will start in 5 months time. A department wants to create a team with a leader now, and give the leader the next 5 months to select appropriate team members as she pleases. Is this possible?

4. Client Z has used the same supplier for the past 5 years. The supplier charges at a discount rate and bills at the end of each phase. Does the model allow the purchase of consumables if there are no key deliverables?

The fourth set of materials comprised a "personal-profile" questionnaire to obtain information about participants' academic qualifications, industrial experience, and modeling experience.

3.3 Participants

Participants in the experiment were 20 individuals working in industry. They did not play an information technology role in their organizations, nor did they have information systems or technology qualifications. In essence, in the experiment they acted as surrogate end users. Demographic data was collected, but it is omitted here for reasons of brevity. All had at least a Bachelor's degree. Twelve had no experience of data models. The remainder had minor experience of one or two modeling techniques like flowcharts or financial models.

3.4 Procedures

Participants were first assigned randomly to one of the two treatments (10 per treatment). They were then run singly through the experiment. When they arrived to undertake the experiment, they were asked to complete a consent form and the demographic survey.

Next they were given the document that explained the UML symbols. Participants were permitted to discuss the symbols with the researchers until they indicated they felt confident with the UML symbols. They retained and could refer to the UML summary throughout the experiment.

When participants indicated they were ready to begin, they were then given either the ontologically sound or ontologically unsound UML class diagram. They retained and could refer to the UML class diagram throughout the experiment. During the experiment, participants were asked to speak aloud. Their verbalizations were tape-recorded. In addition, the times they took to answer each problem-solving question were recorded. Notes were also made based on participant reactions, queries, and approaches to each question. One researcher conducted the experiment, while another took notes and observed the participant's behavior during the experiment.

4 RESULTS

Scores for the individual items on the problem-solving dependent measures were calculated. Statistical analyses were performed on the scores for each dependent measure.

4.1 Data Scoring

The maximum score for each problem-solving question was four. The score was calculated as follows:

1. *Answer*: One mark was given if the answer ("possible" or "not possible") was correct; zero was given if the answer was incorrect.

2. *Tacit knowledge*: A full mark was given if the participant did not use tacit knowledge while attempting the question. A half mark was given if only slight use of tacit knowledge occurred. No marks were given if tacit knowledge was clearly used. A judgment on the marks to be awarded was made based on the explanations that participants provided to support their answers, the notes taken by one of us during the experiment, and the audio recording made during the experiment.

3. *Explanation*: A judgment on the marks to be awarded for this item was made based on participants' explanations, our notes, and the audio recording. Clear explanations (matching sample answers) to support an answer were awarded one mark. Moderately clear explanations (partially matching sample answers) were awarded a half mark. Unclear explanations (not matching sample answers) were awarded zero marks.

4. *Interpretation*: Again, a judgment on the marks to be awarded for this item was made based on participants' explanations, our notes, and the audio recording. Clear interpretations (matching sample answers) of those aspects of the model on which the problem-solving question focused were awarded one mark. Moderately clear interpretations (partially matching sample answers) were awarded a half mark. Unclear interpretations (not matching sample answers) were awarded zero marks.

Two researchers independently scored the problem-solving measures on preformatted scoring sheets. Few differences arose between the two sets of scores. Where they did occur, they were discussed and reconciled.

4.2 Data Analysis

For each group, Table 1 shows the mean *accuracy* scores and standard deviations for each problem-solving question and the problem-solving questions in total. We applied the Bonferroni family adjustment to the level of significance and set our a value at 0.05/11 or 0.0045 to avoid getting false positives. This level is very conservative.

For 4 of the 11 questions, Table 1 also shows that the ontologically sound treatment group outperformed the ontologically unsound treatment group ($p < 0.0045$). For the remaining questions, the treatment means were not statistically different. For the overall means (based on the 11 questions in total), the ontologically sound treatment group outperformed the ontologically unsound treatment group ($p < 0.001$).[2]

For the problem-solving *time* performance measure and the *normalized accuracy* scores, there were no statistically significant differences for any question or for the overall means. In the interests of brevity, the detailed results have not been reported. In summary, we obtained moderate support for our proposition based on the accuracy measure of problem-solving performance.

Table 1. Accuracy Performance on Problem-Solving Questions

Question	**Ontologically Sound (n = 10)** **Mean (Standard Deviation)**	**Ontologically Unsound (n = 10)** **Mean (Standard Deviation)**	**t-Statistic (2-tail Significance)**
1	3.350 (1.248)	2.450 (1.235)	1.621 (.122)
2	2.850 (1.226)	1.650 (1.132)	2.275 .035
3	3.300 (1.085)	1.200 (0.483)	5.590 (.000)
4	3.100 (1.220)	3.200 (1.252)	-0.181 (.858)
5	3.300 (1.033)	1.400 (0.966)	4.249 (.000)
6	3.150 (1.454)	1.200 (0.632)	3.899 (.001)
7	3.000 (1.414)	2.500 (1.130)	0.632 (.535)
8	3.200 (1.033)	2.850 (1.415)	1.299 (.210)
9	3.350 (1.107)	1.300 (1.059)	4.231 (.001)
10	2.750 (1.399)	3.350 (1.107)	-1.063 (.302)
11	3.450 (1.212)	2.600 (1.265)	1.534 (.142)
Total	34.800 (6.571)	23.700 (5.067)	3.912 (.001)

[2]Statistical tests were also undertaken on the four components of the accuracy score. Basically, the results held for the accuracy, explanation, and interpretation components of the scores. There were no differences between the two groups, however, for the tacit-knowledge component of the score.

5 IMPLICATIONS OF THE RESEARCH

Our results have implications for both practice and research. For practice, we have support for our proposition that composites should be modeled explicitly as an entity class and not modeled implicitly as an association class. The evidence indicates that practitioners should be circumspect if they model composites implicitly as relationships. They run the risk they will undermine users' understanding of the real-world phenomena being represented.

For research, our results add strength to a growing body of empirical work that supports the usefulness of ontological theories, especially Bunge's (1977) ontological theory, as a means of predicting the strengths and weaknesses of conceptual modeling grammars and practices (e.g., Bodart et al. 2001; Gemino 1999; Green and Rosemann 2000; Opdahl and Henderson-Sellers 2001; Parsons 1996; Parsons and Wand 2000; Weber 1996). Specifically, ontological theories allow us to pinpoint which features of conceptual modeling grammars and practices are likely to be problematical and to then design empirical research to test our predictions. This approach stands in stark contrast to previous approaches that attempted omnibus feature comparisons or case-study comparisons of different grammars and methods (e.g., Olle et al. 1983). The equivocal results produced using such approaches motivated calls for better theory to guide conceptual modeling research (e.g., Floyd 1986).

Our research also highlights the need for researchers to reflect carefully on the ways in which they measure users' understanding of the phenomena represented by a conceptual model. For instance, Bodart et al. (2001) argue that researchers need to decide whether they wish to test users' *surface* understanding of phenomena or *deep* understanding of phenomena. Free-recall measures, comprehension measures, and perhaps requirements-verification methods might be appropriate to test users' *surface* understanding of phenomena represented by a conceptual model. Stronger tests like problem-solving tests appear to be required, however, to test users' *deep* understanding of the phenomena represented by a conceptual model. The problem-solving tasks used in this study resemble scenarios or use cases that are widely employed in requirements acquisition in practice. Thus, they are highly relevant when measuring users' understanding of conceptual models.

6 LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

The major limitations of our research that we perceive relate to statistical conclusion validity and external validity. In terms of statistical conclusion validity, we would prefer to have a larger sample size. Given the onerous nature of our experiment, however, we could not obtain more business people willing to participate in the experiment without incurring substantial costs. Moreover, many of our key tests (the problem-solving tests) are statistically significant anyway, even with the small sample size.

In terms of external validity, as with most experiments the context of the experimental task is limited in scope and a little artificial. Nonetheless, we believe our task has enough realism that our results should be robust in other settings involving part-whole relations.

Future research work might be pursued in three directions. First, ontological theory can be used to predict the strengths and weaknesses of other conceptual modeling practices. For instance, UML provides various approaches to modeling the dynamics of a domain. Bunge's (1977) theory might be used to predict whether these approaches are likely to facilitate or inhibit users' understanding of the dynamics of a domain that involves composite things. Second, more work needs to be done on developing valid and reliable measures of users' understanding of the semantics of a domain. Our research suggests that problem-solving measures have merit. Nonetheless, more work is needed to determine how measures of understanding might take into account that (1) users create their worlds (Hirschheim et al. 1995) and (2) shared meaning among a cohort of users may or may not exist. Third, alternative methods of having users validate conceptual models as representations of their perceived worlds might be investigated. Our research suggests that methods based on having users solve problems with conceptual models have merit. Nonetheless, a more-systematic articulation of and evaluation of different methods needs to be undertaken.

7 ACKNOWLEDGMENTS

An Australian Research Council Discovery Grant funded this research. We thank Daniel Moody, Steve Hitchman, and Graeme Simsion for helpful comments on our research.

7 REFERENCES

Batra, D., Hoffer, J. A., and Bostrom, R. P. "Comparing Representations with Relational and EER Models," *Communications of the ACM* (33), 1990, pp. 126-139.

Bodart, F., Sim, M., Patel, A., and Weber, R. "Should Optional Properties be Used in Conceptual Modeling? A Theory and Three Empirical Tests," *Information Systems Research* (12:4), 2001, pp. 384-405.

Boehm, B. *Software Engineering Economics.* Upper Saddle River, NJ: Prentice-Hall, 1981.

Bunge, M. *Treatise on Basic Philosophy: Volume 3: Ontology I: The Furniture of the World*. Boston: Reidel, 1977.

Chen, P. P. S. "The Entity-Relationship Model: Toward a Unified View of Data," *ACM Transactions on Database Systems* (1:1), 1976, pp. 9-36.

Elmasri, R., and Navathe, S. B. *Fundamentals of Database Systems* (3rd ed.). Reading, MA: Addison-Wesley, 2000.

Floyd, C. "a Comparative Evaluation of System Development Methods," in T. W. Olle, H. G. Sol, and A. A. Verrijn-Stuart (eds.), *Information System Design Methodologies: Improving the Practice.* Amsterdam: North-Holland, 1986, pp. 19-54.

Gemino, A. *Empirical Methods for Comparing System Analysis Modeling Techniques*, unpublished Ph.D. thesis, University of British Columbia, 1999.

Goldstein, R. C., and Storey, V. "Some Findings on the Intuitiveness of Entity-Relationship Constructs," in F. H. Lochovsky (ed.), *Entity-Relationship Approach to Database Design.* Amsterdam: Elsevier Science Publishers, B. V., North Holland, 1990, pp. 9-23.

Green, P., and Rosemann, M. "Integrated Process Modeling: An Ontological Evaluation," *Information Systems* (25:2), 2000, pp. 73-87.

Halpin, T. A. *Conceptual Schema and Relational Database Design: A Fact Oriented Approach* (2nd ed.). Sydney: Prentice-Hall, 1995.

Hirschheim, R., Klein, H., and Lyytinen, K. *Information Systems Development and Data Modeling: Conceptual Foundations and Philosophical Foundations*. Cambridge, UK: Cambridge University Press, 1995.

Kilov, H., and Ross, J. *Information Modeling: An Object-Oriented Approach.* Upper Saddle River, NJ: Prentice-Hall, 1994.

Mayer, R. E. "Models for Understanding," *Review of Educational Research* (59), 1989, pp. 43-64.

Moody, D. *Dealing with Complexity: A Practical Method for Representing Large Entity-Relationship Models,* unpublished PhD thesis, University of Melbourne, 2001.

Moody, D. L., and Shanks, G. "Improving the Quality of Entity-Relationship Models: An Action Research Programme," *The Australian Computer Journal* (30), 1998, pp. 129-138.

Olle T. W, Sol, H. G., and Tully, C. J. (eds.). *Information System Design Methodologies: A Feature Analysis*. Amsterdam: North-Holland, 1983.

Opdahl, A. L., and Henderson-Sellers, B. "Grounding the OML Metamodel in Ontology," *Journal of Systems and Software* (57), 2001, pp. 119-143.

Opdahl, A. L., Henderson-Sellers, B., and Barbier, F. "Ontological Analysis of Whole–Part Relationships in OO-Models," *Information and Software Technology* (43), 2001, pp. 387-399.

Parsons, J. "An Information Model Based on Classification Theory," *Management Science* (42:10), 1996, pp. 1437-1453.

Parsons, J., and Wand, Y. "Emancipating Instances from the Tyranny of Classes in Information Modeling," *ACM Transactions on Database Systems* (25:2), 2000, pp. 228-268.

Prietula, M. J., and March, S. T. "Form and Substance in Physical Database Design: An Empirical Study," *Information Systems Research* (2), 1991, pp. 287-314.

Rumbaugh, J., Jacobson, I., and Booch, G. *The Unified Modeling Language Reference Manual*. Reading, MA: Addison-Wesley, 1999.

Simons, P. *Parts: A Study in Ontology.* Oxford: Clarendon Press, 1987.

Smith, B. "Mereotopology: A Theory of Parts and Boundaries," *Data & Knowledge Engineering* (20), 1996, pp. 287-303.

Smith, J. M., and Smith, D. C. P. "Database Abstractions: Aggregation," *Communications of the ACM* (20:6), 1977, pp. 405-413.

Storey, V. "Meronymic Relationships," *Journal of Database Administration* (2:3), 1991, pp. .22-35.

Teorey, T. J., Yang, D., and Fry, J. P. "A Logical Design Methodology for Relational Databases Using the Extended Entity-Relationship Model," *ACM Computing Surveys* (18:2), 1986, pp. 197-222.

Varzi, A. C. "Parts, Wholes, and Part-Whole Relations: The Prospects of Mereotopology," *Data & Knowledge Engineering* (20), 1996, pp. 259-286.

Wand, Y., Storey, V., and Weber, R. "An Ontological Analysis of the Relationship Construct in Conceptual Modeling," *ACM Transactions on Database Systems* (24), 1999, pp. 494-528.

Wand, Y., and Weber, R. "Information Systems and Conceptual Modeling: A Research Agenda," *Information Systems Research* (13), 2002, forthcoming.

Wand, Y., and Weber, R. "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars," *Journal of Information Systems* (3), 1993, pp. 217-237.

Weber, R. "Are Attributes Entities? A Study of Database Designers' Memory Structures," *Information Systems Research* (7), 1996, pp. 137-162.

Winston, M. E., Chaffin, R., and Herrman, D. "A Taxonomy of Part-Whole Relations," *Cognitive Science* (11), 1987, pp. 417-444.