

Towards an Enhanced Methodology for Ontological Analyses

Michael Rosemann¹, Peter Green², Marta Indulska²

¹Centre for Information Technology Innovation
Queensland University of Technology
126 Margaret Street, Brisbane QLD 4000
m.rosemann@qut.edu.au

²UQ Business School, The University of Queensland
11 Salisbury St, Ipswich QLD 4305, Australia
{p.green, m.indulska}@business.uq.edu.au

Abstract. The ontological analysis of conceptual modelling techniques is of increasing popularity. Related research did not only explore the ontological deficiencies of classical techniques such as ER or UML, but also business process modelling techniques such as ARIS or even Web services standards such as BPEL4WS. While the selected ontologies are reasonably mature, it is the actual process of an ontological analysis that still lacks rigor. The current procedure leaves significant room for individual interpretations and is one reason for criticism of the entire ontological analysis. This paper proposes a procedural model for the ontological analysis based on the use of meta models, the involvement of more than one coder and metrics. This model is explained with examples from various ontological analyses.

1 Popularity of Ontological Analyses

As techniques for conceptual modelling, enterprise modelling, and business process modelling have proliferated over the years (e.g.,[9]), researchers and practitioners alike have attempted to determine objective bases on which to compare, evaluate, and determine when to use these different techniques (e.g.,[3, 8]). Throughout the 80's, 90's, and into the new millennium however, it has become increasingly apparent to many researchers that a theoretical foundation was needed on which to base the specification for these various modelling techniques.

Wand and Weber [16-20] have investigated the branch of philosophy known as ontology as a foundation for understanding the process in developing an information system. Ontology is a well-established theoretical domain within philosophy dealing with identifying and understanding elements of the real world. The popularity of using ontologies as a basis for the analysis of techniques that purport to assist analysts to develop models that emulate portions of the real world has been growing steadily. The Bunge-Wand-Weber (BWW) ontological models [21], for example, have been applied extensively in the context of the analysis of various modelling techniques. Wand and Weber [16-20] and Weber [21] have applied the BWW representation model

to the “classical” descriptions of entity-relationship (ER) modelling and logical data flow diagramming (LDFD). Weber and Zhang [22] also examined the Nijssen Information Analysis Method (NIAM) using the ontology. Green [5] extended the work of Weber and Zhang [19, 20, 22] and Wand and Weber [19, 20] by analysing various modelling techniques as they have been extended and implemented in upper CASE tools. Furthermore, Parsons and Wand [12] proposed a formal model of objects and they use the ontological models to identify representation-oriented characteristics of objects. Along similar lines, Opdahl and Henderson-Sellers [10] have used the BWW representation model to examine the individual modelling constructs within the OPEN Modelling Language (OML) version 1.1 based on “conventional” object-oriented constructs. Green and Rosemann [6] have extended the analytical work into the area of integrated process modelling based on the techniques presented in Scheer [14]. Most recently, Green *et al.* [4] have extended the use of this evaluative base into the area of enterprise systems interoperability using business process modelling languages like ebXML, BPML, BPEL4WS, and WSCI.

While ontological analyses are frequently utilised, particularly in the area of conceptual modelling technique analysis, the actual process of performing the analysis remains problematic. The current process of ontological analysis is open to the individual interpretations of the researchers who undertake the analysis. Consequently, such analyses are criticised as being subjective, *ad hoc*, and lacking in relevance. There is a need, therefore, for the systematic identification of shortcomings of the current ontological analysis process. The identification of such weaknesses, and their subsequent mitigation, will lead to a more rigorous, objective, and replicable analytical process.

Accordingly, this paper has two main objectives. First, we aim to identify comprehensively the shortcomings in the current practice of ontological analysis. The identification of such shortcomings will provide a basis upon which the practice of ontological analysis can be improved. Second, we want to develop several propositions and methodology extensions that enhance the ontological analysis process by making it more objective and structured.

There are several contributions that this paper aims to make. First, the work presents a detailed analysis of the actual process of performing an ontological evaluation. The presented work identifies eight shortcomings of the current ontological analysis process. Each of the identified shortcomings is classified then as belonging to one of three phases of analysis, *viz.*, input, process, and output. Second, the paper presents recommendations on how each of the shortcomings in the three phases can be overcome. The recommendations, *inter alia*, include an extended methodology for the improvement of the objectivity of the analysis.

The remainder of this paper is structured as follows. The next section identifies eight current shortcomings of ontological analyses that are classified with respect to the three phases of analysis, *viz.*, input, process and output. The third section provides recommendations concerning how to overcome the identified shortcomings in each of the three phases. The final section provides a brief summary of this work and outlines future research in this area.

2 Shortcomings of Current Ontological Analyses

An ontological analysis is in principle the evaluation of a selected modelling grammar from the viewpoint of a pre-defined and well-established ontology. The current focus of ontological analyses is on the bi-directional comparison of ontological constructs with the elements of the modelling grammar that is under analysis. Weber (1997) clarifies two major situations that may occur when a grammar is analysed according to an ontology. After a particular grammar has been analysed, predictions on the modelling strengths and weaknesses of the grammar can be made according to whether some or any of the following situations arise out of the analysis.

1. *Ontological Incompleteness (or Construct Deficit)* exists unless there is at least one grammatical construct for each ontological construct.
2. *Ontological Clarity* is determined by the extent to which the grammar does not exhibit one or more of the following deficiencies:
 - *Construct Overload*
 - *Construct Redundancy*
 - *Construct Excess*

Though this type of ontological analyses is widely established, it still has a range of shortcomings. These shortcomings can be categorised into the three main phases of an ontological analysis, *i.e.* preparation of the input data, the process of conducting the analysis, and the evaluation and interpretation of the results.

The first two identified shortcomings refer to the quality of the input data.

2.1 Lack of Understandability

Most of the ontologies that are currently used for analysis of modelling grammars have been specified in formal languages. While such a formalisation is beneficial for a complete and precise specification of the ontology, it is not naturally a very intuitive specification. An ontology that is not clear and intuitive can lead to misinterpretations as the involved stakeholders have problems with the specifications. Furthermore, it forms a hurdle for the application of the ontology as it requires a deep understanding of the formal language in which it is specified.

2.2 Lack of Comparability

The specification of an ontology requires typically a formal syntax, which allows the precise specification of the elements and relationships of the ontology. Such specifications are required, but not necessarily intuitive. Consequently, textual descriptions of the ontology in 'plain English' often extend the formal specification.

However, even if an ontology is specified in an intuitive and understandable language, the actual comparison with the selected modelling grammar remains a problem. Unless the ontology and the grammar are specified in the *same* language, it will be up to the coder to 'mentally convert' the two specifications into each other, which

adds a subjective element to the analysis. Different languages can also lead to different levels of detail and further complicate the analysis.

The additional three shortcomings identified below are related to the process of the ontological analysis and refer to what should be analysed, how it should be analysed, as well as who should conduct the analysis.

2.3 Lack of Completeness

The first decision that has to be made in the process of an ontological analysis is on the scope and depth of the analysis. Even if most ontologies have been discussed for many decades they still undergo modifications and extensions. It is up to the researcher to clearly specify the selected version of the ontology and the scope and level of detail of the analysis. In our work in the area of Web Services, for example, it was often not clear what constructs form the core of the standard. Two researchers, who conducted independent analyses of the same Web Services standard, selected consequently a different number of constructs.

Moreover, many ontological analyses solely focus on the constructs of the ontology and the constructs of the grammar but do not sufficiently consider the relationships between these constructs. The difficulty in clearly specifying the boundaries of the analysis as well as the limited consideration of relationships between the ontological constructs lead to a lack of completeness.

2.4 Lack of Guidance

After the scope and the level of detail of the analysis have been specified, it is typically up to the coder to decide on the procedure of the analysis, i.e. in what sequence will the ontological constructs and relationships be analysed? Currently, there are hardly any recommendations on where to start the analysis. This lack of procedural clarity underlies most analyses and has two consequences. First, a novice analyst lacks guidance in the process of conducting the ontological evaluation. Second, the procedure of the analysis can potentially have an impact on the results of the analysis.

2.5 Lack of Objectivity

An ontological analysis of a grammar requires not only detailed knowledge of the selected ontology and grammar but also a good understanding of the languages in which the ontology and the grammar are specified. This requirement explains why most analyses are carried out by single researchers as opposed to research teams. Consequently, these analyses are based on the individual interpretations of the involved researcher, which adds significant subjectivity to the results. This problem is further compounded by the fact that, unlike other qualitative research projects, ontological analyses typically do not include attempts to further validate the results.

The five shortcomings identified above have a common flavour in that they heavily depend on the researcher conducting the ontological evaluation. Three further short-

comings have been identified, *viz.*, lack of result representation, lack of result classification and lack of relevance. These shortcomings are detailed below and refer to the outcomes of the analysis

2.6 Lack of Adequate Result Representation

The results of a complete ontological analysis, i.e. representation and interpretation mapping, are typically summarised in two tables. These tables list all ontological constructs (first table) and all grammatical constructs (second table) and the corresponding constructs of the other meta model. Such tables can become quite lengthy and are typically not sorted in any particular order. They do not provide any insights into the importance of identified deficiencies and they also do not cluster the findings.

2.7 Lack of Result Classification

As indicated above, it is common practice to derive ontological deficiencies based on a comparison of the constructs in the ontology and the grammar. Such identified deficiencies are a typical starting point for derivation of propositions and then hypotheses. In general, the ontological analysis does not make any statements regarding the relative importance of these findings in comparison with each other. This lack of a more detailed statement regarding the significance of a potential shortcoming makes it difficult to judge quickly the outcomes of the results of two different sets of analyses, *e.g.*, an ontological analysis of ARIS in comparison with an ontological analysis of UML.

2.8 Lack of Relevance

Finally, the results of an ontological analysis should be perceived as relevant by the related stakeholders. However, if an ontological analysis leads, for example, to the outcome that Entity Relationship models do not support the description of behaviour, then it is not surprising that the IS community develops a rather critical opinion. It seems that an ontological analysis has to consider the purpose of the grammar as well as the background of the modeller who is applying this grammar. The application of a high-level and generic ontology does not consider this individual context and there is a danger that the outcomes can be perceived as trivial.

3 A Reference Methodology for Conducting Ontological Analyses

The above identified shortcomings motivated the development of an enhanced methodology for ontological analyses. The main purpose of this methodology is to increase the rigour, the overall objectivity and the level of detail of the analysis. The

proposed methodology for ontological analyses is structured in three phases, *viz.*, input, process and output.

3.1 Input

The formal specification of ontologies, together with the differences in the languages used to specify the ontologies and the grammars under analysis, have been classified as issues pertaining to the lack of understandability and comparability.

In order to overcome these shortcomings, it is proposed to convert the ontology as well as the selected modelling grammar to meta models using the same language (e.g. ER Models or UML Class Diagram). This conversion facilitates a pattern-matching approach towards the ontological analyses of completeness and clarity of a grammar. As a first step we converted, for example, the Bunge-Wand-Weber ontology into an ER-based meta model. This meta model includes 50 entity types and 92 relationship types. It has clusters such as *system*, *property* or *class/kind*. Such a meta model explains, in a language familiar to the Information Systems (IS) community, the core constructs of the ontology. It also highlights the underlying focus of the ontology. The obtained meta model can now be used for a variety of ontological analyses. Moreover, it allows a critical review of the BWW model by a wider community. The approach, however, is not without its limitations. Commonly used modelling techniques such as ER or UML are often widely accepted, however, they have not been designed for the purposes of meta modelling. Thus, they lack occasionally the required expressiveness.

While an ER-based meta model helps to overcome issues related to the understandability of an ontology, a corresponding meta model of the analysed grammar is required to deal with the lack of comparability issue. Many popular modelling techniques (e.g. ARIS or UML, and also interoperability standards such as ebXML) are already specified in meta models using ER-notations or UML Class Diagrams. If the meta models for the ontology and the modelling technique are specified in the same language, the ontological analyses turns into a comparison of two conceptual models. It also becomes immediately obvious whether the paradigm of the analysed grammar differs from the ontology or not. In the case of ARIS or many web services standards, for example, the meta models are centred around *functions* or *activities* instead of being centred around *things*.

3.2 Process

Issues related to the process of conducting an ontological analysis have been described as lack of completeness, lack of guidance and lack of objectivity.

Based on the assumption that corresponding meta models for the ontology and the analysed grammar are available, it is possible to clearly specify the scope of an analysis using those meta models. Such a selection of clusters, entity types and relationship types would define all elements that are to be perceived of relevance for a complete analysis. An analysis of an ER-based notation, for example, could be focused on the BWW clusters *thing*, *system* and *property* and could exclude the more behavioural-

oriented clusters *event* and *state*. Such boundaries of an analysis could be easily visualised in the meta model and would provide a clear description of the comprehensiveness of the analysis.

The existence of two corresponding meta models and a clear definition of the scope of the analysis is a necessary but not a sufficient criteria for a well-guided process. Further guidelines are required regarding the starting point of such a process and the actual sequence of activities. Based on our experiences, we recommend starting with the representation mapping, i.e. selecting the meta model of the ontology and subsequently identifying the corresponding elements in the modelling grammar. The first construct to be analysed should be the most central entity type, i.e. in the case of the BWW models the entity type *thing*. Our previous work provides a strong argument that this analysis if followed by a cluster-by-cluster approach. Starting with the core constructs in a cluster, this allows a more structured and focused analysis of the completeness of a modelling grammar. The representation mapping is followed by an analysis of the clarity, i.e. the interpretation mapping. In this case the meta model of the grammar under analysis is the starting point. The general procedure is similar. A main advantage of a cluster-based analysis is that the structure of the two meta models provides valuable input for the ontological analysis. An example is the analysis of generalisation-specialisation relationships in the meta model of the grammar. We propose to ontologically classify the super-type first and then to inherit this ontological classification to all sub-types. This streamlines the process of the analyses and increases the consistency.

The lack of objectivity issue, on the other hand, frequently stems from the analysis being performed by a single researcher. The situation results in an analysis that is almost certainly biased by the researcher's background as well as their interpretation of the grammar specification. In order to improve the validity of the analysis, a research methodology can be adopted that undertakes individual analyses of a grammar by at least two members of a research team, followed by consensus as to the final analysis by the entire team of researchers. The methodology consists of three steps:

Step 1: Using the specification of the grammar in question, at least two researchers separately read the specification and interpret, select and map the ontological constructs to candidate grammatical constructs.

Step 2: The researchers involved in Step 1 of the methodology, meet to discuss and defend their interpretations of the representation modelling analysis. This meeting leads to an agreed second draft version of the analysis. The overlap in the selection of the constructs and in the actual ontological analysis can be quantified by various figures that are used in content analysis and other more qualitative research.

Step 3: The second draft version of the analysis for each of the interoperability candidate standards is used as a basis for defense and discussion in a meeting involving the entire research team. The outcome of this meeting forms the final analysis of the grammar in question.

Such a methodology was employed in a project that applied the BWW representation model analysis to a number of the leading potential Web Service standards, *viz.*, ebXML, BPML, BPEL4WS and WSCI. The project team was composed of four researchers and the standards were analysed in the order: ebXML → BPML → BPEL4WS → WSCI. Two researchers were involved in steps 1 and 2 of the method-

ology. This stage was followed by a meeting of the entire team in order to discuss the mapping and arrive at the final analysis.

The adoption of such a methodology is seen to have greatly improved the objectiveness of the carried out analyses.

3.3 Output

The three main shortcomings related to the outcome of an ontological analysis have been characterised as the lack of adequate result representation, lack of result classification and the lack of relevance.

The meta models, which have been used as input for the ontological analyses, are an appropriate medium to visualise the outcomes of the entire analysis process. In our work on the analysis of ARIS, we derived a meta model of the BWW model that highlighted all constructs of the ontology that do not have a corresponding construct in the grammar under analysis, *i.e.* we visualised incompleteness in the model using simple colour coding. In a similar way, we derived three ARIS meta models that highlighted excess, overload and redundancy in ARIS. Such models form a very intuitive way of representing the identified ontological shortcomings. The underlying clustering of the models also helps to quickly comprehend the main areas of shortcomings.

At present time, the process of an ontological analysis results in the identification of ontological incompleteness and ontological clarity through the identification of missing, overloaded or redundant grammatical constructs. While the end result identifies such problems, it fails to account for their relative importance. For example, *thing* is one of the fundamental constructs of the BWW model. The lack of mapping for the construct should, therefore, be considered more important than the lack of mapping for the *well-defined event* construct for example. There is a need for the development of a scoring model that enables the calculation of the 'goodness' of a grammar with respect to the ontology. In such a scoring model, core constructs would therefore have high weightings whereas less important constructs would attract lower values of weightings. Following an ontological analysis of a particular grammar, the weighting of all missing constructs would be calculated to arrive at one value that generally reflects the outcome of the analysis. The scores would be aggregated across the ontology and modelling grammar. They also would be calculated separately for completeness, excess, overload and redundancy. Furthermore, they could be aggregated per cluster, which allows a more differentiated view on the particular strengths of a modelling grammar. Though the consolidated score of such an evaluation should not be overrated, it provides better insights into the characteristics of the ontological deficiencies and provides a first rating of the significance and importance of the identified shortcomings.

Apart from the lack of result classification that is addressed by the scoring model, another problem with the outcome of the analyses has been the perceived lack of relevance. Since most modelling grammars focus on modelling a sub-set of the phenomena that occurs in the real world, it would follow that not *all* constructs of an ontology are necessary in order to analyse such a grammar. If the full ontology is used in the analysis, the result may identify potential problems that would not, in

reality, occur. Further, there may also be a need for specialisation of some of the ontological constructs in order to enhance analysis of a grammar pertaining to a particular domain.

Indeed, the outcomes of the ontological analyses of different modelling grammars to date appear to support the need for a focused ontology, which consists of different subsets of the ontological constructs for different domains. The analyses of the examined grammars consistently show that the constructs *conceivable state space*, *conceivable event space* and *lawful event space*, for example, have no representative constructs in the grammars. Such missing constructs, if identified to be unnecessary for the particular domain, can be ignored leading to a simpler analysis that does not consider phenomena that are deemed to be outside of the scope of the domain.

4 Summary and Future Work

There has been a marked increase in the popularity of the application of ontologies for the purposes of modelling grammar analysis. For example, a literature review identified more than 25 papers that applied the Bunge-Wand-Weber ontology for the analysis of modelling grammars such as ER [16, 19, 20], OMT, UML [2, 11, 15], Petri-Nets, ARIS [6, 7, 13] or Web Services standards such as ebXML, BPEL4WS, BPML or WSCI [1, 23]. In general, selected ontologies and their interpretations, from an Information Systems viewpoint, are reasonably advanced. However, the actual process of conducting an ontological analysis is still rather pre-mature.

In total, eight shortcomings of the current process of ontological analysis have been identified and categorised into issues related to the input, process and output of the analysis. This paper proposed to further enhance the current process of ontological analyses. The objectives of such a methodology are

- to provide guidance for researchers who are interested in conducting ontological analyses
- to add rigour to the entire process and reduce the dependence on the subjective interpretations of the involved researcher, and
- to overall increase the credibility of the ontological analysis.

Examples from our ontological analyses of ARIS and various Web Services standards have been used to exemplify this methodology. As a consequence, we hope that the presented more rigorous process will increase the overall acceptance of using ontologies for the analysis, comparison and engineering of various grammars.

References

1. van der Aalst, W.M.P., Dumas, M., A.H.M. ter Hofstede and Wohed, P.: Pattern Based Analysis of BPML (and WSCI). Queensland University of Technology: Brisbane (2002)
2. Burton-Jones, A., Meso, P.: How good are these UML diagrams?: An empirical test of the Wand and Weber good decomposition model. 23rd International Conference on Information Systems. Barcelona (2002)

3. Gorla, N., Pu, H.-C., Rom, W.O.: Evaluation of process tools in systems analysis. *Information and Software Technology*, 37(2) (1995) 119-126
4. Green, P., Rosemann, M., Indulska, M., Manning, C.: *Candidate Interoperability Standards: An Ontological Overlap Analysis*. The University of Queensland (2003)
5. Green, P.F.: Use of Information Systems Analysis and Design (ISAD) Grammars in Combination in Upper CASE Tools - An Ontological Evaluation. 2nd CaiSE/IFIP8.1 International Workshop on the Evaluation of Modeling Methods in Systems Analysis and Design. Barcelona (1997)
6. Green, P.F., Rosemann, M.: Integrated Process Modelling: An Ontological Evaluation. *Information Systems*, 25(2) (2000) 73-87
7. Green, P.F., Rosemann, M.: Perceived ontological weaknesses of process modeling techniques: Further evidence. in 10th European Conference on Information Systems. 2002. Poland.
8. Karam, G.M. and Casselman, R.S.: A cataloging framework for software development methods. *IEEE Computer*, 1993. February: p. 34-46.
9. Olle, T.W., Hagemstein, J., Macdonald, I.G., Rolland, C., Sol, H.G., Assche, F.J.M.V., Verrijn-Stuart, A.A.: *Information Systems Methodologies: A Framework for Understanding*. Wokingham: Addison-Wesley (1991)
10. Opdahl, A.L., Henderson-Sellers, B.: Grounding the OML metamodel in ontology. *Journal of Systems and Software*, 57(2) (2001) 119-143
11. Opdahl, A.L., Henderson-Sellers, B.: Ontological evaluation of the UML using the Bunge-Wand-Weber Model. *Software and Systems Modeling*, 1(1) (2002) 43-67
12. Parsons, J., Wand, W.: Using objects in systems analysis. *Communications of the ACM*, 40(12) (1997) 104-110
13. Rosemann, M., Green, P.: Developing a meta model for the Bunge-Wand-Weber ontological constructs. *Information Systems*, 27(2) (2002) 75-91
14. Scheer, A.-W.: *ARIS - Business Process Modeling*. 3 ed., Heidelberg: Springer: Berlin (2000)
15. Shanks, G., Tansley, E., Nuredini, J., Tobin, D., Weber, R.: Representing part-whole relationships in conceptual modelling: An empirical evaluation. 23rd International Conference on Information Systems. Barcelona (2002)
16. Wand, Y., Weber, R.: An ontological evaluation of systems analysis and design methods, in *Information System Concepts: An In-depth Analysis*, E.D. Falkenberg and P. Lindgreen, Editors. North-Holland (1989) 79-107
17. Wand, Y., Weber, R.: Mario Bunge's Ontology as a formal foundation for information systems concepts, in *Studies on Mario Bunge's Treatise*, P. Weingartner and G.J.W. Dorn, Editors. Rodopi: Atlanta (1990) 123-149
18. Wand, Y., Weber, R.: An ontological model of an information system. *IEEE Transactions on Software Engineering*, 16(11) (1990) 1281-1291
19. Wand, Y., Weber, R.: On the ontological expressiveness of information systems analysis and design grammars. *Journal of Information Systems*, 3(4) (1993) 217-237
20. Wand, Y., Weber, R.: On the deep structure of information systems. *Information Systems Journal*, 5 (1995) 203-223
21. Weber, R.: *Ontological Foundations of Information Systems*. Coopers & Lybrand accounting research methodology monograph ; no. 4. Melbourne, Vic.: Coopers & Lybrand and the Accounting Association of Australia and New Zealand (1997)
22. Weber, R., Zhang, Y.: An analytical evaluation of NIAM's grammar for conceptual schema diagrams. *Information Systems Journal*, 6(2) (1996) 147-170
23. Wohed, P., W.M.P. van der Aalst, Dumas, M., A.H.M. ter Hofstede: *Pattern Based Analysis of BPEL4WS*. Queensland University of Technology: Brisbane (2002)