Ontological Clarity and Comprehension in Health Data Models

Sophie Cockroft
University of Queensland, sophie@business.uq.edu.au

Follow this and additional works at: https://aisel.aisnet.org/amcis2005

Recommended Citation
https://aisel.aisnet.org/amcis2005/437

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2005 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
ABSTRACT

Conceptual modeling forms an important part of systems analysis. If this is done incorrectly or incompletely, there can be serious implications for the resultant system, specifically in terms of rework and useability. One approach to improving the conceptual modelling process is to evaluate how well the model represents reality. Emergence of the Bunge-Wand-Weber (BWW) ontological model introduced a platform to classify and compare the grammar of conceptual modelling languages.

This work applies the BWW theory to a real world example in the health arena. The general practice computing group data model was developed using the Barker Entity Relationship Modelling technique. We describe an experiment, grounded in ontological theory, which evaluates how well the GPCG data model is understood by domain experts. The results show that with the exception of the use of entities to represent events, the raw model is better understood by domain experts.

Keywords
Conceptual modelling, Information systems development, Ontology, Subtypes

INTRODUCTION

The aim of this paper is to present the results of a study concerning an empirical test of how well the GPCG data model is understood. The experiment involves a two-group test. The between group factor is type of representation with one group reviewing the raw model and one group reviewing a model recast according to ontological principles.

A body of literature is emerging in the field of health data modelling. Specific models under development include the Good Electronic Health Record (Bird, Goodchild and Tun, 2003; Bird, 2000; Bird, 2003) which has recently been incorporated into the OpenEHR initiative (Blobel, 2004). The Health level 7 Reference Information Model (HL7 RIM) is also still an active area for research (Dolin, Alschuler, Beebe, Biron, Boyer, Essin, Kimber, Lincoln and Mattison, 2001) as is the Clinical Document Architecture (CDA) which is an XML version of HL7.

Much work is also being undertaken in the evaluation and analysis of conceptual modelling formalisms using the BWW framework. This work is detailed in the following section

This study is distinguished from previous studies in that it uses ontological analysis to review an instance of a data model rather than the grammar itself. It builds on the approach of Kruse (1999) who sought to evaluate an instance of ER modelling used in practice, namely the HL7 model (Dolin et al., 2001). Through the use of the BWW model Kruse found that it was possible to analyse the scripts and understand them from an ontological viewpoint. A subset of the HL7 Model (Accounting and Administrative functions) was re-modelled from the ontological point of view, identifying areas of representational deficiency, and several advantages of this recast model were recognised. Evaluation of a concrete example demonstrated the merits of the BWW model and the need for information systems grammar to be rooted in a theory of ontology.

The approach used by Kruse, and followed here, was alluded to a recent paper by Shanks, Tansley et al (2003) who suggested that theories of ontology can provide an indication of how modelling grammars should be used to model phenomena in the focal domain.

The paper proceeds as follows; first some background is provided into the GPCG data model, and the representation model of the BWW. This leads to the theoretical foundations of the current work and particular issues emerging from an initial data gathering study of the GPCG are identified. The research method is the elucidated followed by the initial results and a discussion of these results with issues for further investigation flagged.
BACKGROUND

The General Practice Computing Group (GPCG) Model

The GPCG model has been developed in Australia under a government grant to improve general practice information systems, and to create an industry standard. It was developed in cognisance of a number of other models including the reference information model of HL7 (Dolin et al., 2001) and the Good Electronic Health Record. The full model covers issues pertaining to the patient, schedules, service providers, decision support, programs/trials and clinical knowledge. For the purposes of this study this was narrowed down to the patient domain. The full model has 68 core entity types and 125 subtypes. The patient domain consists of 11 Entities and 53 subtypes. The service provider domain has 17 entities and 22 subtypes. In this study all the core entities of the patient domain are included. Since some core entities are only related through their subtypes the relevant subtypes are also included.

The model uses a the Barker Entity Relationship modelling technique (Barker, 1990) distinguished in this case by its use of nested subtypes. A review of this technique is given in Halpin (2000).

This modelling technique is used in CASE tools from vendors such as the oracle corporation, and was adopted by the systems analysts on the GPCG project, Simsion Bowles & Associates (2000). The model is publicly available.

Overview of the Bunge-Wand-Weber (BWW) model

The BWW model (Wand and Weber, 1989) has been widely used to analyse and evaluate conceptual modelling grammars. The fundamental premise behind the BWW model is that information systems are a representation of real world systems. One aspect of the BWW model is the representational model. A grammar’s completeness and clarity is reduced due to the following:

- **Completeness** is undermined in cases of a Construct Deficit; this means the grammar does not have a construct corresponding to an identified real world construct
- **Clarity** is challenged where there are instances of the following
  - Construct redundancy; More than one grammatical construct maps on to an ontological construct
  - Construct excess; a grammatical construct may have no corresponding ontological construct

Table 1 Summarizes recent work in the area of evaluating conceptual modelling grammars, with a particular emphasis on those studies that have included empirical tests. In designing the questionnaire used in this study the GPCG data model was examined for instances of the deficits outlined above

<table>
<thead>
<tr>
<th>Construct</th>
<th>Citation</th>
<th>Key Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes and Entities</td>
<td>(Weber, 1996)</td>
<td>Discovered that sophisticated users of the NIAM modelling technique which does not explicitly distinguish between entities and attributes, still distinguished between them in a recall exercise.</td>
</tr>
<tr>
<td>Relationship</td>
<td>(Wand, Storey and Weber, 1999)</td>
<td>Put forward rules to resolve ambiguities that exist for users of conceptual models concerning whether an association should be shown via a relationship, an entity or an attribute domain.</td>
</tr>
<tr>
<td>Optional properties</td>
<td>(Bodart, Patel, Sim and Weber, 2001)</td>
<td>Optional properties should be used where the goal is surface understanding, they should not be used where deep understanding is required because they undermine users’ abilities to grasp important domain semantics.</td>
</tr>
<tr>
<td>Whole-Part Relationships</td>
<td>(Opdahl and Henderson-Sellers, 2001), (Shanks, Tansley and Weber, 2004)</td>
<td>Opdahl and Henderson-Sellers carried out a formal analysis of whole part relationships leading to a distinction between primary, consequential, secondary and dependant characteristics of Whole-Part (WP) relationships. Shanks et al put forward a case that composites (e.g. WP relationships) should be represented as entities not relationships or associations. Both acknowledged the plethora of unresolved issues in the use of WP relationships in conceptual modelling.</td>
</tr>
</tbody>
</table>
Relationship with Attributes (Burton-Jones and Weber, 1999) Burton Jones and Weber argued that the construct of a relationship with attributes produces unclear representations of a domain.

Structural Constraints (Siau, Wand and Benbasat, 1997) Showed that users of conceptual modelling would take notice of structural constraints (e.g. cardinality and modality) even when they were in contradiction of the real world situation.

Properties and Things (Shanks, Nuredini, Tobin, Moody and Weber, 2003) Things and their properties should be clearly distinguished in conceptual models.

Classes and Instances (Parsons, 1996; Parsons and Wand, 2000) Conceptual models should clearly distinguish between classes and instances.

<table>
<thead>
<tr>
<th>Table 1 Previous work on validating the use of conceptual modelling constructs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>THEORY AND PROPOSITIONS</strong></td>
</tr>
</tbody>
</table>

Parts of the GPCG model have been recast according to key principles described in specifically this meant examining how the entities relationships and attributes are used within the model and redefining them with respect to the relevant rules specified in that paper.

- Rule 1: Things are represented only as instances. Instances should represent only things. Things are fundamental constructs that should not be used to represent other ontological concepts. Wand et al (1999) highlighted associative entities specifically for their violation of this rule. An example in the GPCG model is the service professional booking, which represents an associiative entity between appointment and diary element, and by extension timeslot. The notion of an associative entity is well accepted in the conceptual modeling literature, such literature often recommends that entities or objects (i.e. things) be used to represent constructs other than things, in the case of the associative entity it is known as the “event remembered” and it connects two object classes where the relationship is of cardinality M:N.

- Rule 2: Both simple and composite things should be represented using the same construct so relationships should not be used to represent composite thing. This is avoided in the GPCG model by drawing entities within entities, according to the Barker ER nested subtype notation instead of using the is-a construct. There is no instance of where this rule is violated.

- Rule 3: A class or kind of a thing is defined in terms of a given set of intrinsic attributes and relationships; that is, intrinsic attributes and mutual attributes. Prescription item is quite problematic in this regard since as the model stands it models both a class and instance of a thing (i.e. drugs).

- Rule 4: An aggregate type/class must have properties in addition to those of its component types/classes. In this model Person was initially recast as a generalization of patient, and it does have attributes in addition to Patient, for example a person can also be a Doctor. However the original intention of the model was to use these two entities separately since at the top of the hierarchy is GP client which is a generalization of Patient and Group Patient and is defined as a candidate entity.

- Rule 5: All attributes and relationships in a class represent properties of things in the class.

- Rule 6: Null Attributes have no meaning. A few of these emerge in the model in such areas as a patient may have zero or many care plans and a health problem may be referenced to zero or many care plans.

- Rule 7: The same construct should be used to represent a binary relationship and a higher order relationship. No higher order relationships were modeled in the original GPCG model.

- Beyond these seven rules, a clear example of construct overload in the GPCG data model is the use of entities to represent events.

Based on the rules expressed above, and their related examples from the raw model, the following three propositions were developed.

Proposition 1: Conceptual schema diagrams that use things to represent instances and classes in the same model will be harder to understand than those that do not.

Proposition 2: Conceptual schema diagrams that do not use things to represent events will assist the user to undertake tasks that require a surface level understanding of a domain better that those that do.

Proposition 3: Conceptual schema diagrams that do not use nested subtypes will assist the user to undertake tasks that require a surface level understanding of a domain better than diagrams that do.
OVERVIEW OF THE EXPERIMENT

The aim of the study was to test the ease of interpretation between two models an experimental approach, and particularly a two group test, is an appropriate means of measuring such difference. The participants in the study were medical students, who were in their clinical years and had considerable experience in practice; they were mature students having undergraduate degrees in cogent disciplines. Part of their training involves a GP placement, but none were familiar with the GPCG data model. Participants were recruited via their faculty office by email and notices placed around the faculty. Each was paid AUD$25 to encourage them to participate in the experiment and attempt to complete it satisfactorily.

Recasting the data Model

The models were recast to remove construct overload, construct redundancy and construct excess. In creating the questionnaire the areas for investigation focused on incidence of overload identified at the end of the previous section, the problem of associative entities within the appointment-health professional-facilities relationship (rule 1), the prescription item class/instance (rule 3) and how encounter data is drawn from the GP service event. It is postulated that construct overload is the key shortcoming of the Barker notation. The models and the questionnaire were validated by means of a pre-test conducted in the business school, the results of which appear in (Cockcroft and Rowles, 2003). In addition the recast model was reviewed by an expert in the BWW model.

Instances and classes in the same model

As foreshadowed above, the prescription item in the raw model represents both a class and an instance of a thing, i.e. drugs. Thus a questionnaire item was developed relating to this i.e., does prescription item refer to a class of items e.g. aspirin, or a particular instance of an item e.g. packet of aspirin no: 345532 just dispensed to that patient?

The use of entities to represent events

Entities representing events were removed from the raw model and included in a separated state transition diagram in the recast model. Examples of these are the GP service event, and Communications that occur between health professionals. It has been shown (Allen and March, 2003) that an event-oriented graphical representation leads to more efficient query formulation, probably due to the role of events and causality in human problem solving.

Associative entities, optional properties and optional relations

A problem was encountered in recasting the associative entities within the appointment-health professional-facilities relationship. The original model does not show properties, they are available in an attached spreadsheet of descriptions. The Service Professional Diary Entry, and the Facility Diary Entry are both associative entities. Two real world things, an appointment, and a Facility/Health Service Professional Diary Element participate in a many to many association with each other. One of the constructs we would seek to avoid is optionality, which can occur under two conditions;

1. where a thing may or may not possess an attribute or
2. where a thing may or may not participate in a relationship with another thing.

These conditions are described in detail in Bowen et al (2004). For ontological clarity in the first example, one would create a subclass representing those things that possess a property and a subclass representing those that do not. In order to do this with the GPCG it would first be necessary to add attributes to the raw model, this option was ruled out. In the second example to derive a model with greater ontological clarity one subclass is created which represents those things that possess the properties of the thing being modelled and another subclass that represents those things that represent the optional aspect. This was attempted in the recast model. The effect was essentially to remove the nested subtypes.

Materials and Procedure

According to the classification given in Bodart, Patel et al. (2001) the experiment was a comprehension test, the type of meaning transfer evaluated was mostly surface level, the independent variable was type of representation, and the dependent variables were response accuracy and response time. The use of a comprehension test was in order to measure performance only with respect to the semantics in the model. A recent analysis of prior experimental research in conceptual modeling (Parsons and Cole, 2003) has illustrated that this type of test is more appropriate in situations where participants have to answer questions using only a script and model.
A questionnaire was developed with which to gather data on the areas of representational deficiency outlined above. The full questionnaire, raw and recast models are given in Appendix 1. The experimental materials are designed to explore how well the model was understood, by means of a two-group test. The between-group factor is "type of representation" with one group reviewing the raw model, and one group reviewing the recast model. The dependent variable, performance, is evaluated by assessing responses to a short multiple choice questionnaire. The make up of the review group was as shown in Table 2. Participants were given up to 30 minutes to complete the task with the times being recorded as they completed the questionnaire and left the room. The distribution of times between the two groups was equivalent. Overall Score data was normally distributed.

<table>
<thead>
<tr>
<th>Pre Medical Training</th>
<th>Re-Cast</th>
<th>Raw</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health related</td>
<td>8</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Science</td>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Years of Medical/Related Education</th>
<th>&lt;=3</th>
<th>4-6</th>
<th>&gt;6</th>
<th>Not Stated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average time taken to complete the test</td>
<td>24.62</td>
<td>24.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions relating to Instances and Classes</td>
<td>56.67</td>
<td>74.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions relating to Events</td>
<td>72.73</td>
<td>72.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions relating to nested subtypes</td>
<td>78.48</td>
<td>82.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Numbers and background of participants in each review group

RESULTS

The preliminary results are presented in Table 3 and a T test is presented in Table 4. The T-test demonstrates that the difference in overall mean scores between the two groups is statistically significant. Further Table 3 reveals that participants scored better on the raw model except for those responses that relate to events. Thus propositions 1 and 3 are rejected. A number of possible reasons for this are put forward below.
### Group Statistics

<table>
<thead>
<tr>
<th>Model Type</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score/34</td>
<td>0</td>
<td>11</td>
<td>22.45</td>
<td>1.440</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13</td>
<td>25.54</td>
<td>2.295</td>
</tr>
</tbody>
</table>

### Independent Samples Test

<table>
<thead>
<tr>
<th>Score/34</th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>1.708</td>
<td>.205</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>-4.002</td>
<td>20.448</td>
</tr>
</tbody>
</table>

Table 4 T-test for statistical significance of difference in scores between groups (raw and ontologically recast)

The results suggest that the raw model was easier to understand than the recast model, further investigation is indicated to ascertain the exact cause of this. Likely reasons would be: First, the use of domain experts they could use their knowledge to answer the comprehension tests as opposed to the scripts/models. Second, the fact that recasting the model caused an increase in size and necessitated participants in the recast study viewing more than one sheet.

### LIMITATIONS OF THE RESEARCH

It would be possible to rebuild the model with optional relationships and attributes removed instead of measuring comprehension of nested subtypes, although that approach needs careful consideration given the nature of the raw model. Given the models were reviewed by domain experts, the opportunity for tacit knowledge to enter into a comprehension task is a possibility. The sample is small, and whilst this did not reduce statistical conclusion validity, it means that the study has limited generalizability.

### FURTHER WORK

Since this study was conducted, work has progressed in the area of experimental design for evaluating conceptual models, notably Parsons and Cole (2005). It would be useful to carry out a larger study adhering to the guidelines put forward by them regarding experimental design and subject selection. With respect to the use of domain experts, this deficiency might represent an opportunity in terms of re-running the experiment with a set of non domain experts with the purpose of pooling the results and exploring the effects of domain knowledge. The study might also include controlling for domain knowledge versus familiarity or modeling knowledge.

### CONCLUSION

The experiment found that a focal domain data model when recast according to ontological principles was actually harder to understand. It is suggested that this is in part a consequence of the increased model size a phenomenon recently noted by other researchers (Bowen et al, 2004). The research represents one of the few attempts to apply ontological principles to a real world model and for this reason is a contribution to knowledge. It also represents a contribution to the field of experimental design in this area, and some of the pitfalls encountered in the rigorous application of ontological principles.

### REFERENCES

APPENDIX 1

QUESTIONNAIRE

Please attempt all questions

1 Focus on the **GP Service Event**. When a patient visits a health professional, a number events can occur. Please answer the following questions:

a) When a patient visits a GP, does triage have to occur before they are dispensed any drugs or undertake any procedure?  
Y/N

b) According to the model:

(i) Could a consultation with a GP result in another consultation ?  
Y/N
(ii) Could a consultation with a GP result in no tests, procedures or dispensing events?  
Y/N
(iii) Could a consultation result in three more consultations, one for a test, one for a procedure and one at which a prescription was written  
Y/N
(iv) Could a consultation result in a test, procedure and prescription all at the same time?  
Y/N
(v) Could an encounter test result in a procedure?  
Y/N

c) Is any data stored about the patient as a result of GP service events?  
Y/N

d) Does a GP review take place

(i) Prior to an encounter?  
□
(ii) After an encounter?  
□
(iii) Independently of an encounter?  
□
(iv) Any of the above?  
□

2 You will notice an entity called **GP encounter data**, which has subclasses Subjective, Objective, Assessment and Plan (SOAP)

a) Does the model indicate that SOAP data is associated with:

(i) A GP service event  
Y/N
(ii) A GP encounter  
Y/N
(iii) A GP consultation  
Y/N

b) Supposing a patient goes through a GP consultation, and is sent off for a test as a result, can you discern the following from the model? (if you can’t please select X)

(i) Are all types of data (SOAP) stored as a result of this test.  
Y/N/X
(ii) Is it possible for no data to be stored as a result of this test  
Y/N/X
3 Focus on Communication

a) Does a communication between parties involved in the care of a patient have to concern a specific health problem?
   Y/N

b) A health summary communication (HSC) is drawn from elements of the patients profile as well as clinical observations. It could be a discharge summary, transfer of care, or general referral. It may have requests attached – such as review in three weeks.

(i) According to the model, can an HSC be related to more than one patient? Y/N
(ii) From the model, can you detect any need to update a patient record as a result of this type of communication? Y/N

c) A service order is a type of communication; it describes a service (pathology, radiology etc) and instructions to enable the provider to perform the service for the patient

(i) From the model is it possible for such a communication to recommend many services? Y/N
(ii) Does each service order item have to be associated with a result? Y/N
(iii) Can a service order item have more than one result? Y/N

d) From the model, do service order results relate in a 1:1 form with service order items. For example if a service order included a blood glucose level and a blood haemoglobin level, could the service order result include:

(i) Glucose results
(ii) Haemoglobin results
(iii) Glucose and haemoglobin results
(iv) Other unexpected results
(v) All of the above
(vi) Impossible to tell

(e) A prescription is a type of communication; it is issued to a pharmacist and includes treatments and medications to be dispensed to the patient. A prescription item is part of the prescription; it describes a specific pharmaceutical product to be dispensed to the patient. A prescription item authorization indicates that the prescriber is approved by the authorities to prescribe the item in question.

(i) Is the prescription item a type of communication? Y/N
(ii) Can a prescription item be dispensed without an authorization? Y/N
(iii) Can an item appear on more than one prescription? Y/N
(iv) Can a prescription be for no items at all? Y/N

f) Mr Fang the dentist issues Mr Ho with a prescription for viagra. This is not on the schedule of drugs permitted to be prescribed by dentists and subsidised by the government, however a recent ruling has determined that if it is prescribed privately, and the patient pays for it, it is allowed. Does the model allow for this situation?
   Y/N
g) If Mr Fang includes a number of permitted drugs on the prescription (e.g. painkillers, antibiotics) does a separate authorisation need to appear on the prescription for each drug, according to the model?
Y/N

4 Finally turn your attention to **Diary Elements and Diary Entries**

a) A facility commitment is when a facility (such as a room) is tied up for other reasons than an appointment, for example for maintenance. According to the diagram, do I need to make an _appointment_ to tie the interview room up for cleaning on Thursday afternoons (this is a generic time slot, ie no specific date associated with it)?
Y/N

b) Dr Smith wants to see a patient in the operating theatre on Friday from 10-11 and Mr Jones the anaesthetist needs to be present. Does the model allow for an appointment to have many health service professionals?
Y/N

c) Suppose Dr Jones has booked to be on holiday that morning (ie she is a health service professional with a commitment), would the model allow us to book her for an appointment at this time?
Y/N

If No, which particular feature of the model prevents this?

__________________________________________________________________

d) A new doctor arrives in the practice, and a new consultation room is built for them. Which part of the model allows new schedulable time slots to be created for these new resources, if any?