December 2003

Ontological Evaluation of Health Models: Some Early Findings

Sophie Cockcroft
University of Queensland

Scott Rowles
University of Queensland

Ontological Evaluation of Health Models: Some Early Findings

Sophie Cockcroft[a], Scott Rowles[a]

[a]UQ Business School, University of Queensland
Blair Drive, St Lucia, Queensland 4072
s.cockcroft@business.uq.edu.au
s.rowles@business.uq.edu.au

Abstract

Incomplete or incorrect systems analysis can result in expensive rework at a later stage in system development. It is in the interests of all parties involved in a system development endeavour to get the data model right before development begins.

The General Practice Computing Group (GPCG) data model and Core Data Set is an emerging industry standard for General Practice systems. It describes and documents clinical activities in general practice, facilitates the exchange of information between general practice and the broader health system as well as providing a foundation for the development of general practice computing applications.

Emergence of the Bunge-Wand-Weber (BWW) ontological model introduced a platform to classify and compare the grammar of conceptual modelling languages. It established a systematic and theoretical basis for the evaluation of grammars. This approach has been adapted in an earlier study, which evaluated a specific instance of a conceptual model rather than the model itself. The model evaluated was in the health domain. The work presented here follows the approach of the earlier study but takes it one step further by gathering empirical evidence to test the hypothesis that ontologically correct models more readily support different users conceptions of a domain.

Grounding a model according to ontological principles is an important step in establishing how well it represents reality. This is something that should be undertaken alongside the more ambitious aim of realising the vision of a life-long, portable, fully integrated, health record, and implementing it on a global scale.

The work uses the Bunge-Wand-Weber (BWW) model of ontological correctness. This approach is being used increasingly for the validation of conceptual models. This study differs from earlier ones in that it uses the BWW model to identify problems in an instance of a data model rather than the modelling grammar itself. This paper sets out the theoretical foundations for evaluating this model and presents the results of a pilot study

Keywords

Health Informatics, Ontology, Data Model Comprehension

Introduction

The aim of this paper is to present the results of some initial data gathering concerning an empirical test of the comprehensibility of the GPCG data model. The experiment involves a two-group test. The between-group factor will be "type of representation" with one group reviewing the raw model, and one group reviewing a model re-cast according to ontological principles

It is possible to resolve ambiguities that exist in current data modelling practice and rules relating to this resolution can enrich the capacity of an entity- relationship diagram, or other grammar, to capture knowledge about an application domain (Wand and Weber, 1989). In addition to evaluating grammars used in data modelling, the broader study aims to empirically test whether the more semantically correct models leading from such an evaluation are more understandable and, by extension, a more accurate representation of user requirements.

Much work in the area of health data modelling is currently underway in Australia and internationally. Specific models include the Good Electronic Health Record (GEHR) (Beale, 1999, Beale, 2000, Bird, 2003) which has recently been incorporated in the OpenEHR initiative (OpenEHR, 2003). Health Level 7 (HL7) (HL7, 2000), the Patient Record Architecture (PRA) which is an XML version of HL7, and the General Practice Data Model & Core Data Set Project (GPCG, 2000)

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

This study differs from others in that it uses ontological evaluation to identify flaws in a model in terms of their representation of real world phenomena. A study on HL7 has already been carried out (Kruse, 1999) This work was interesting because instead of using the Bunge-Wand-Weber (BWW) model to identify problems within the modelling languages themselves, it sought to evaluate an instance of ER modelling used in practice, namely the HL7 model. Through the use of the BWW model the Kruse study 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 work proposed here will follow a similar methodology to that of Kruse (1999).

The paper proceeds as follows. In the following section some background to the GPCG data model is given. An overview of the representation model of the BWW model is also presented with a summary of experimental work to date from the literature. This leads into the theoretical foundations of the current work. The research methodology is then elucidated, followed by some initial results and a discussion of these results with pointers for further work in the area.

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. It was developed in cognisance of a number of other models including the reference information model of HL7 (HL7, 2000) and the Good Electronic Health Record (GEHR, 2000). 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 entities 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 (see Figure 1). A review of this technique is given in Halpin (2000).

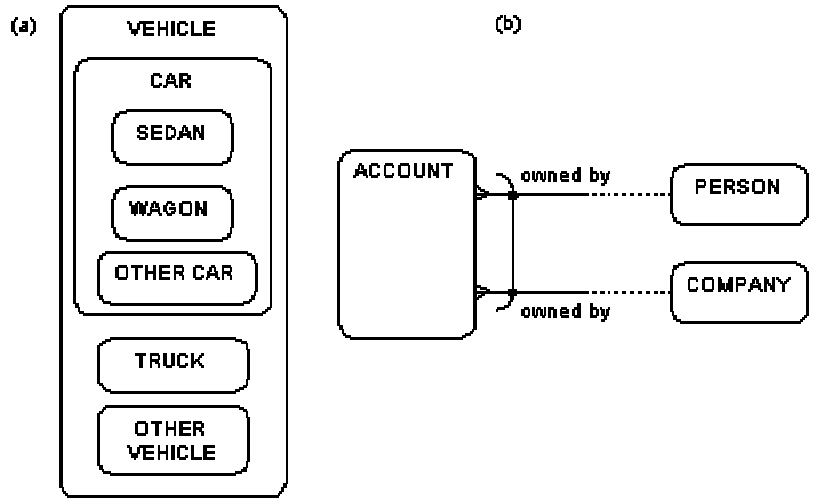


Figure 1 (a) Non-exhaustive, exclusive subtypes in Barker ER and an exclusive-or constraint in the Barker ER notation.

This modeling technique has 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 segment of the model used in this study is given in Appendix 1.

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 overload; A grammatical construct maps onto more than one ontological construct

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 Summarises 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 examples of the above representational deficiencies.

Construct	**Citation**	**Key Finding**
Attributes and Entities	(Weber, 1996)	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.
Relationship	(Wand et al., 1999)	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.
Optional properties	(Bodart et al., 2001)	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
Whole-Part Relationships	(Opdahl and Henderson-Sellers, 2001), (Weber, 2002)	Opdahl et al carried out a formal analysis of whole part relationships leading to a distinction between primary, consequential, secondary and dependant characteristics of WP relationships. Weber 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
Relationship with Attributes	(Burton-Jones and Weber, 2000)	Burton Jones and Weber argued that the construct of a relationship with attributes produces unclear representations of a domain
Structural Constraints	(Siau et al., 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

Table 1. Previous experimental work on validating the use of conceptual modelling constructs

Theoretical foundations

Parts of the GPCG model have been recast according to key principles described in (Wand et al., 1999) 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. This has interesting implications for hitherto well-accepted notion of associative entities. An example of this is the service professional booking, which represents an associative entity between appointment and diary element, and by extension timeslot.

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 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 generalisation 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 generalisation 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 none or many care plan and a health problem may be referenced to none 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 modelled 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.

Research Methodology

The models were recast to remove construct overload, construct redundancy and construct excess.

In creating the questionnaire the areas for investigation focussed 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.

A questionnaire was developed with which to gather data on the areas of representational deficiency outlined above. Initially the instrument was administered to postgraduate students in the business school, the purpose of this approach was twofold. First, we could pilot test the

raw instrument with a group of respondents. Second, we could gather some initial data from which we could gain insight into the shortcomings of the model. In the longer term it is planned to run the experiment on domain experts in the health field.

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 dependant variable, performance, is evaluated by assessing responses to a short multiple choice questionnaire, although in this pilot study the results of participants performance was less important than their feedback on the materials.

The participants in this pilot survey were selected because they were post-graduate students, with knowledge of research design who could provide valuable feedback. A cross section of IS and Non-IS postgraduates was deliberately selected in order to identify which aspects of the model were hard to understand for people not trained in conceptual modelling.

The make up of the review group was as shown in Table 1

	Some Formal ER training	**Non-Expert**
Ontologically Re-cast Model	2	2
Raw Model	2	2

Table 2 Numbers and background of participants in each review group

Results

The initial finding in reviewing the experimental materials was that the ontologically correct version was in fact harder to understand than the raw version, see Table 3.

	Some Formal ER training	**Non-Expert**
Ontologically Re-cast Model	59.6%	45.1%
Raw Model	72.5%	67.7%

Table 3 Average scored for participants in each review group

In comments received from reviewers, two sources of confusion occurred. First with respect to the use of entities to represent events, this was interpreted as a series of mandatory events something akin to a flow chart, not optional subtypes of an event. Second positioning of the optionality/modality of the relationships was critical, in particular the relationship between appointment and facility booking was misinterpreted as being between appointment and Facility Diary Entry.

Discussion of Results

It seems that in an attempt to reduce the ambiguity of using entities to represent events, some richness that existed in the raw model has been lost. A major problem with the Barker notation, identified by (Halpin, 2000) is that is does not depict overlapping subtypes. The use of slightly different notations for the two parts of the study may introduce some noise and needs to be carefully considered before proceeding with the full experiment. Another issue arising from the use of this notation in the raw model is the use of the exclusive-or constraint represented by an arc. In Figure 1 (b) each account is owned by a company or a person, not

both. This representation is apt to cause confusion because viewed on its own account owned by person would appear to be mandatory, but the arc introduces some new semantics changing the meaning to show that the two relationships are disjunctively mandatory. This confusion could also be compounded by the type of effect noted by (Siau et al., 1997) described in Table 1. Despite these two shortcomings it does appear to be at least as easy to understand as the re-cast model. The use of a slightly different graphical style could have, in itself, altered participants' understanding. This phenomenon was explored in (Nordbotten and Crosby, 1999) and (Agarwal et al., 1999). Further, measuring understanding is a difficult task given the varied backgrounds from which participants are likely to come. It has been suggested, however, that task based tests are the best approach in this regard (Weber, 2002). One downside of this approach is the tendency for participants to rely on tacit knowledge rather than the actual model to obtain their answers. The fact that users of data models may resort to domain knowledge to resolve semantic ambiguities is an issue often cited in the literature on ontological modelling where experiments of this kind are used. The phenomenon was first described in the context of human factors. See for example Ashcraft (1989)

Conclusion and Further work

This is a preliminary study, it was designed to assist in the development of a data-gathering instrument. Before proceeding with the full study, it is necessary to review the two models to ascertain if any richness has been lost in the recasting process. Key findings from the review process described are

The limitations of the notation used in the raw model, do not seem to hinder its ability to be understood.

More attention should be paid to the background of the participants in particular the ability of those within the medical domain to use tacit knowledge

In a larger study the effects of data modelling ability or straightforward intellectual capacity should be controlled for.

In further work, the instrument described here will be refined. This will then be presented to domain experts in the health area. An expanded sample will enable better inferences to be made about the propositions of this research, and statistical analysis to be carried out.

References

Agarwal, R., De, P. and Sinha, A. P. (1999) *Comprehending Object and Process Models: An Empirical Study, IEEE Transactions on Software Engineering,* **25,** 541-556.

Ashcraft (1989) *Human Memory and Cognition,* Foresman & Co.

Barker, R. (1990) *CASE Method : entity relationship modelling,* Addison Wesley, Wokingham, England.

Beale, T. (1999) *The GEHR object model architecture,* www.gehr.org, Viewed 20/10/2000

Beale, T. (2000) *GEHR System Architectures,* www.gehr.org, Viewed 20/11/2001

Bird, L. J., Goodchild, A., Tun, Z (2003) *The Good Electronic Health Record in Theory and in Practice*,2003 http://titanium.dstc.edu.au/papers/JRPIT-2002.pdf, Viewed 21/2/2003

Bodart, F., Patel, A., M, S. and Weber, R. (2001) *Should Optional Properties be used in Conceptual Modelling? A Theory and Three Empirical Tests, Information Systems Research,* **12,** 384-405.

Burton-Jones, A. and Weber, R. (2000) In *Twentieth Annual International Conference on Information Systems* (Ed, De, P.) Charlotte, 214-228.

GEHR (2000) *The Good Electronic Health Record,*2000 www.gehr.org, Viewed 22/2/2003

GPCG (2000) *General Practice Computing Group: GP data model and core data set project,* http://www.health.gov.au/hsdd/gp/imit.htm, Viewed 01/12/02

Halpin, T. (2000) *Entity Relationship Modeling from an ORM Perspective: Part 2, Journal of Conceptual Modelling,* **2000,** 1-5.

HL7 (2000) *Health Level 7,* www.hl7.org, Viewed 05/09/00

Kruse, G. (1999) *An ontological analysis and evaluation of the Health Level 7 Data Model, Honours,* University of Queensland

Nordbotten, J. C. and Crosby, M. E. (1999) *The effect of graphic style on data model interpretation, Information Systems Journal,* **9,** 139-155.

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

OpenEHR (2003) *OpenEHR Foundation,* OpenEHR,2003 http://www.gehr.org/openEHR/, Viewed 18 February 2003

Siau, K., Wand, Y. and Benbasat, I. (1997) *The relative importance of structural constraints and surface semantics in information modeling, Information Systems,* **22,** 155-170.

Simsion Bowles & Associates (2000) *General Practice Data Model & Core Data Set Project,* **Version 1.0**.

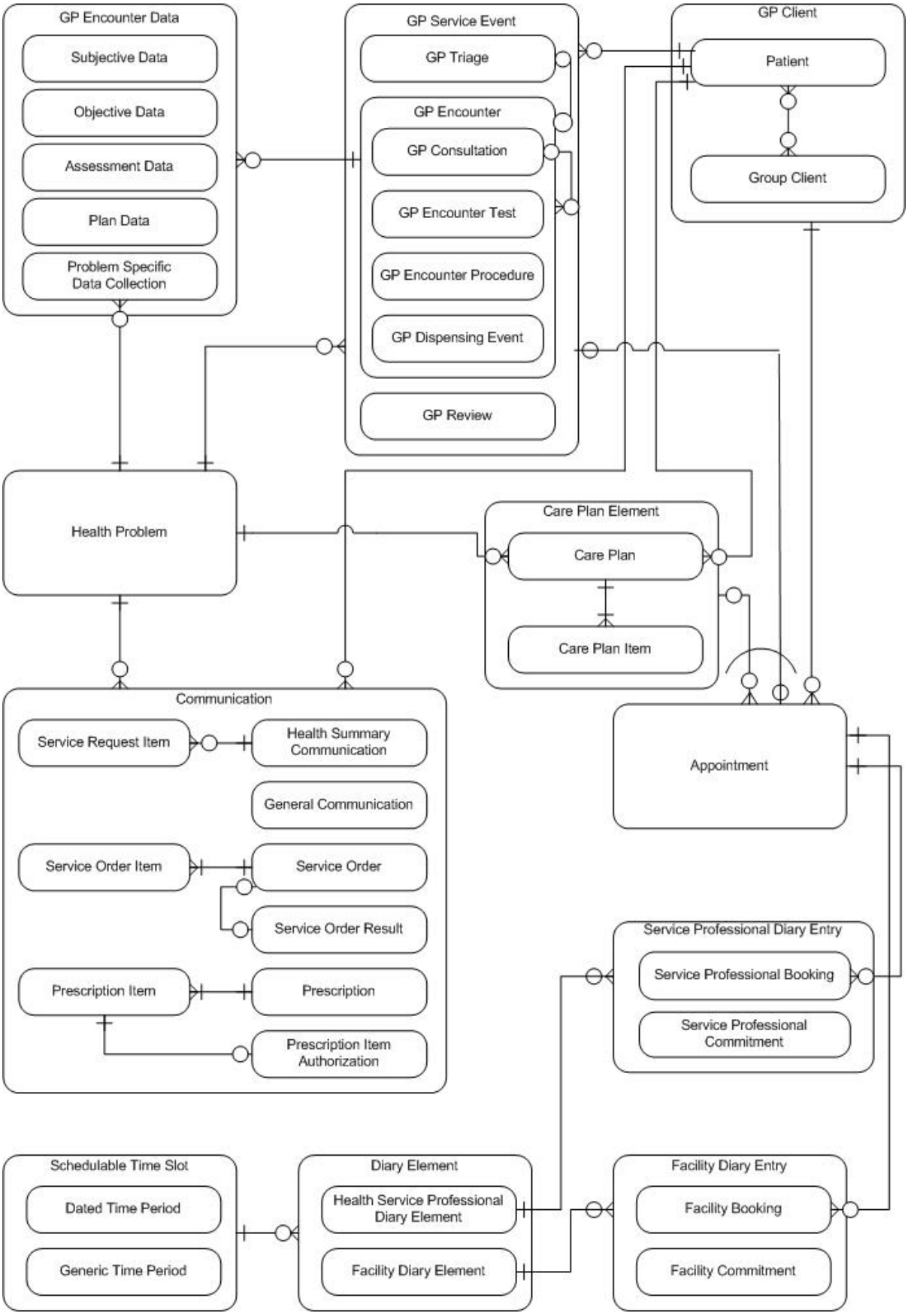
Wand, Y., Storey, V. and Weber, R. (1999) *An ontological analysis of the relationship construct in conceptual modelling, ACM transactions on database systems,* **24,** 494-528.

Wand, Y. and Weber, R. (1989) In *Information system concepts: An in depth analysis*(Eds, Falenberg, E. D. and Lindgreen, P.) North Holland, pp. 79-109.

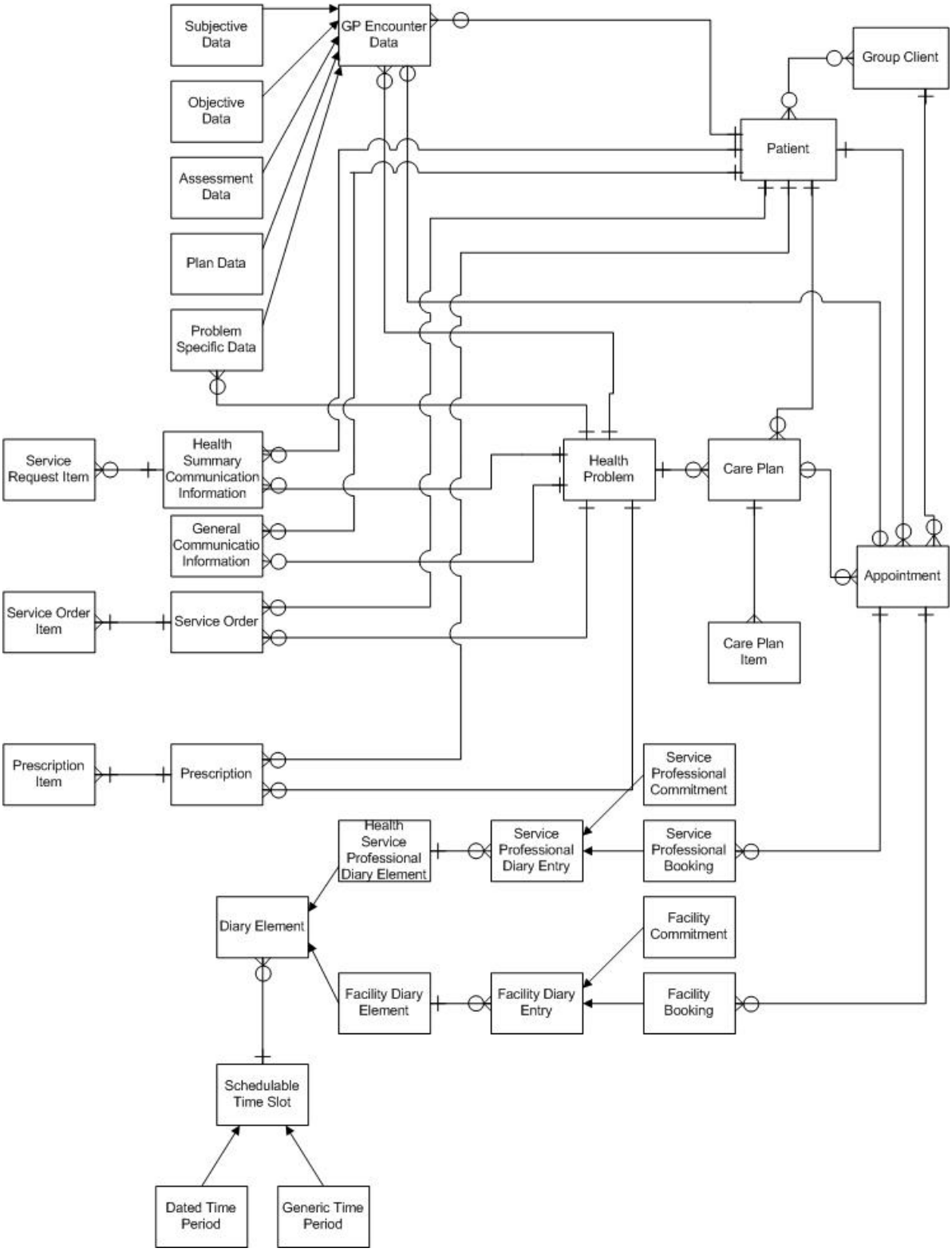
Weber, R. (1996) *Are Attributes Entities? A Study of Database Designers' Memory Structures, Information Systems Research,* **7,** 137-162.

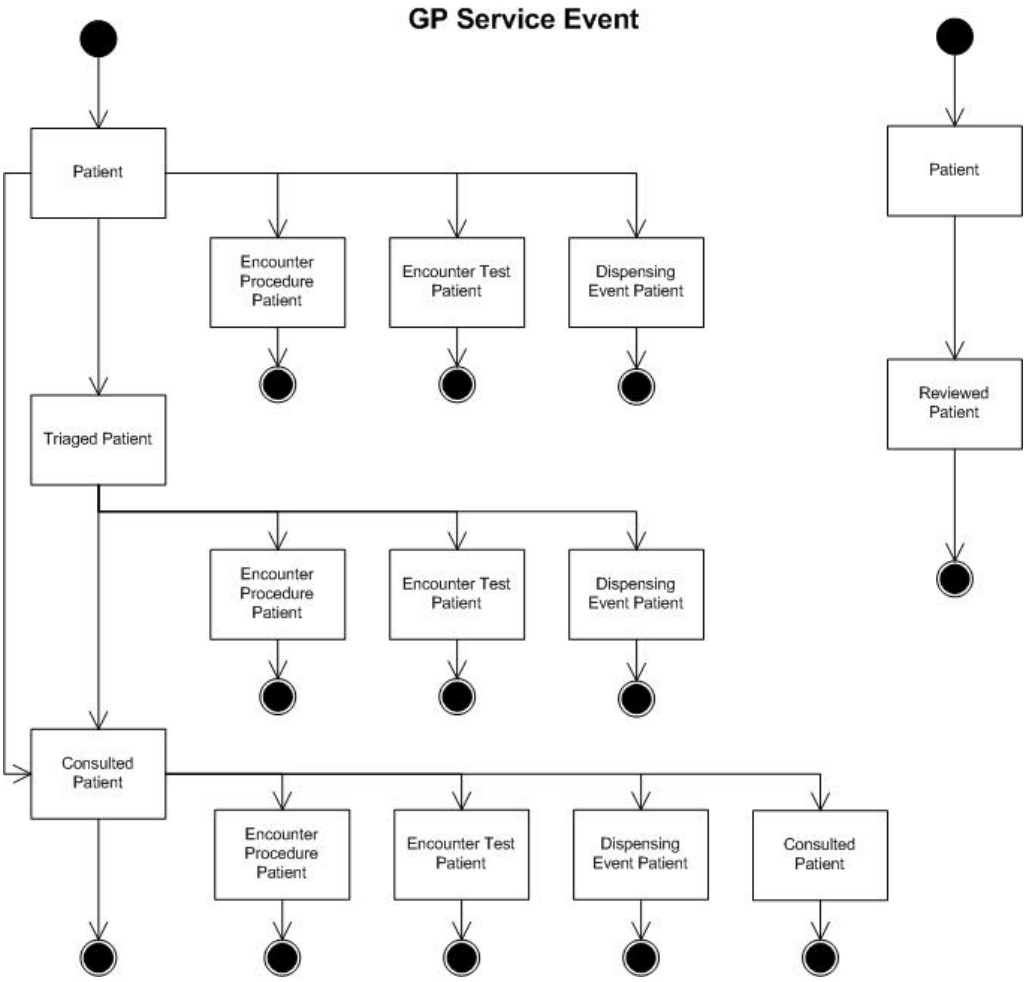
Weber, R. (2002) *Representing Part-Whole relations in conceptual modelling: an empirical evaluation,* University of Queensland, UQ Business School - KPMG Workshop Series.,

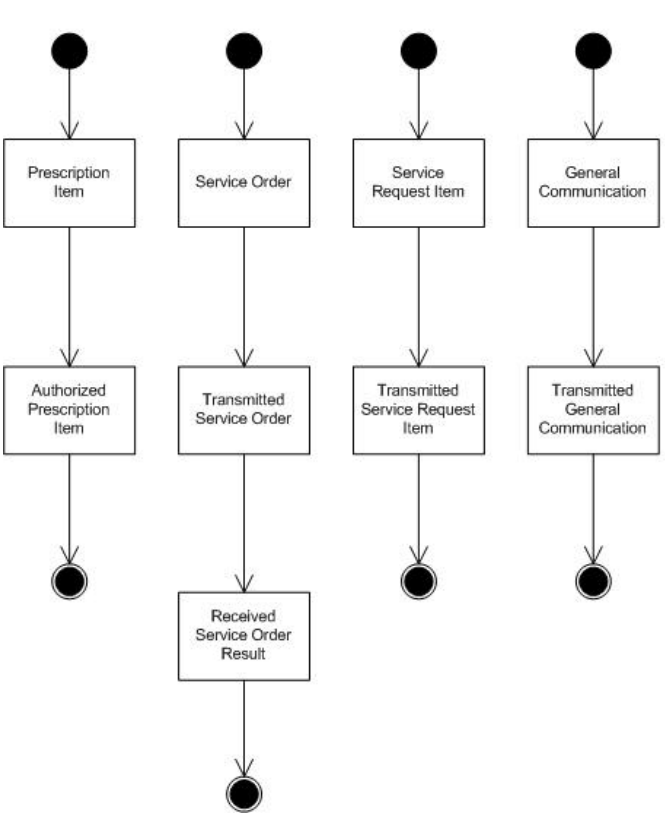
Appendix 1 Raw Data Model Segment



Appendix 2 Recast Model




GP Service Event
Patient
Encounter Procedure Patient
Encounter Test Patient
Dispensing Event Patient
Triaged Patient
Encounter Procedure Patient
Encounter Test Patient
Dispensing Event Patient
Consulted Patient
Encounter Procedure Patient
Encounter Test Patient
Dispensing Event Patient
Consulted Patient
Patient
Reviewed Patient


Communication
Prescription Item
Service Order
Service Request Item
General Communication
Authorized Prescription Item
Transmitted Service Order
Transmitted Service Request Item
Transmitted General Communication
Received Service Order Result

Appendix 3 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 authorisation 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 authorisation? 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 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?

_____