

Developing and evaluating a logical ontology for mediating between nursing intervention terminology systems

N.R. Hardiker

Salford Health Informatics Research Environment, University of Salford, Greater Manchester,
UK

Address for correspondence:

Nicholas R. Hardiker RN PhD

Senior Research Fellow,

Salford Health Informatics Research Environment,

Brian Blatchford Building,

The University of Salford,

Greater Manchester, M6 6PU,

UK.

Tel: +44 161 295 7013

E-mail: n.r.hardiker@salford.ac.uk

Summary

Objectives Several researchers have proposed the use of logical ontologies as ‘reference terminologies’. However, there are a number of unresolved issues. This article describes the development of a logical ontology for nursing interventions and presents the results of evaluation.

Methods Initially this study involved the development in GRAIL of two separate experimental ontologies: an ontology based on the textual content of informal definitions for nursing interventions drawn from the Nursing Interventions Classification; and an ontology based on labels for the same nursing interventions. Following initial bench-testing, the ontology based on labels was selected for extension (to accommodate also nursing intervention components of the Home Health Care Classification System and the Omaha System), for further testing and for external evaluation.

Results A hierarchy of nursing interventions generated automatically from the experimental ontology based on informal definitions contained only 3 hierarchical relationships, compared to 214 for the initial ontology based on labels. For the final extended ontology based on labels, the generated hierarchy contained the three source terminology systems in entirety - there were a total of 2861 hierarchical relationships. While the results of comparative bench testing of the final ontology were favourable, the results of external evaluation were mixed and showed little agreement between reviewers.

Conclusion This study suggests that while a logical ontology based on labels might be a useful tool for mediating between nursing intervention terminology systems, a formative consensus type development methodology might improve the approach by helping to harmonise ideological differences that may exist across the nursing profession.

Keywords: Nursing, Terminology, Classification

1. Introduction

A number of researchers have proposed the use of logical ontologies, symbolic systems for representing concepts and their interrelationships, as mechanisms for mediating between health care terminology systems [1, 2]. Terminology system developers, standards organisations, and health informatics researchers have argued the need for such tools in the quest to support comparability and interchange of data and interoperability between health care applications [3, 4]. Substantial progress has been made. However, within nursing in particular certain important fundamental issues around development and evaluation have yet to be resolved. In order to inform the ongoing debate, the remainder of this article describes the development of an experimental ontology for nursing interventions. It describes some of the key modelling decisions taken and it presents and interprets the results of evaluation.

1.1 The proliferation of nursing terminology systems

In 1994, through the Nursing Information and Data Set Evaluation Center (NIDSECSM) the American Nurses Association (ANA) had recognised 4 terminology systems [5]: the North American Nursing Diagnosis Association classification of nursing diagnoses (NANDA) [6]; the Nursing Interventions Classification (NIC) [7]; the Omaha System (OMAHA) [8]; and the Home Health Care Classification system (HHCC) [9]. Since that time, the number of specialised nursing terminology systems has increased [10]. This proliferation is not necessarily an indication of the inadequacy of any particular terminology system. Rather, the author believes it reflects fundamental limitations with the underlying approach - until relatively recently the only available ‘technologies’ for structuring terminology systems were simple enumerative or combinatorial approaches. Practical experience shows that both enumerative systems and combinatorial systems must be tuned to particular purposes if they are to be useful [11, 12].

Of course there may be overlap concerning scope and coverage, but even where terminology systems intersect there may be a number of lexical and ontological differences *e.g.* different word derivations and different levels of granularity. These differences prevent us from ‘seeing’ overlap without human intervention *i.e.* we are not able directly to make use of the similarities and resolve the differences between diverse nursing terminology systems. Thus some form of

mediation is required in order to facilitate the comparison and interchange of heterogeneous patient care data.

1.2 The GALEN approach

The GALEN approach to mediation as used within this study was borne out of a programme of research and development into enabling technologies that could form the basis for the next generation of clinical information systems [13]. In common with other formal compositional approaches, at the heart of GALEN is a semantically valid model of clinical terminology (*i.e.* an ontology). The ontology is represented in the GALEN Representation and Integration Language (GRAIL) [14]. The ontology consists of entities that are related to one another by attributes to form composite entities. For example, the entity ‘Treatments and Procedures’ might be related to ‘Substance Use’ by an ‘involves’ attribute to form a composite entity to represent ‘Substance Use Treatment’. Within GRAIL there are mechanisms for imposing compositional constraints, for recognising and removing redundancy and for classifying automatically composite entities potentially along multiple axes. For example, as we know that ‘Overdose’ is a specialisation of the more general ‘Substance use’ we can use this information to classify automatically the concept ‘Substance Use Treatment: Overdose’ as a ‘Substance Use Treatment’.

As part of the ontology development process the GALEN approach advocates use of a language for developing simpler intermediate representations, thereby by-passing any need for training in the GRAIL formalism [15]. The intermediate representation language consists of a set of descriptors (corresponding to entities) and a set of semantic links (corresponding to attributes) which define the content of the representation, and a small set of simple constraints to determine the syntax for the intermediate representations. For example the NIC concept ‘Behavior Modification’ might be represented as:

modification

INVOLVES behavior

Intermediate representations are expanded automatically or semi-automatically into more complex GRAIL expressions using the GALEN mapping tool (Tigger). These expressions are then presented to a software system, the Terminology Server [16], for automatic classification within the ontology.

2. Method

Initially this study involved the development and validation of two separate experimental ontologies to determine which approach would be most useful:

- an ontology based on the **textual content of informal definitions** for nursing interventions drawn from the second edition of NIC [17]; and
- an ontology based on **labels** or rubrics for the same nursing interventions [2].

2.1 *The Nursing Interventions Classification*

NIC is a terminology system for describing exclusively the treatments that nurses perform. It is intended for use in all settings and in all specialties. The second edition of NIC contains 433 nursing interventions, each with a label or rubric *e.g.* ‘Analgesic Administration’, an informal definition *e.g.* ‘Use of pharmacological agents to reduce or eliminate pain’, a list of activities that a nurse does to carry out the intervention *e.g.* ‘Check history for drug allergies’, a non-hierarchical code *e.g.* ‘2210’ and a short list of background readings. Although no hierarchical relationships exist between nursing interventions themselves, each intervention is located within a simple taxonomic structure consisting of 30 classes and 7 domains (47 interventions belong to more than one class). NIC was selected as the initial focus for this study as, of the commonly reported nursing intervention terminology systems, it contains the greatest number of pre co-ordinated nursing interventions and thus would provide greater scale.

2.2 *Development of the experimental ontologies*

The focus of the first experiment was on **informal definitions** for nursing interventions drawn from NIC. As shown in the example in Section 2.1 informal definitions within NIC are written in a discursive style, and use a comparatively rich vocabulary. The development of intermediate representations within this first experiment had three phases:

1. An initial modelling activity to develop rapidly a preliminary set of intermediate representations using a relatively literal manual translation process. In order to sustain progress, no attempts were made initially to represent unusual structures. Intermediate representations were not coerced to fit any particular patterns and new descriptors and links were added as necessary.

2. A comparative analysis of word sets derived both from the sources and from the set of initial intermediate representations to determine any omissions and transformations that had been made in the translation from sources to intermediate representations. These included, but were not limited to, transformations concerning word derivations, spelling and prepositions.
3. A more formal process of normalisation to apply consistently to the intermediate representations the transformations identified in the previous phase in an attempt to resolve inconsistencies.

Within the second experiment the focus was on nursing intervention **labels**, again drawn from NIC. Labels within NIC contrast with informal definitions in that they are written in a semi-formal and relatively consistent style. For example, they are made up of 500 individual words (as compared to 1388 in informal definitions), with no prepositions, no possessives and no relative pronouns. Within this second experiment, the development of intermediate representations for labels also had three phases: modelling, analysis and normalisation. However, no formal process of *post hoc* normalisation was needed as consistency between intermediate representations was much higher.

In transforming intermediate representations into GRAIL in both experiments: there was a one-to-one mapping between descriptors and entities and between links and attributes; similar techniques were used to derive the initial hierarchies of elementary GRAIL entities; and compositional constraints were derived and implemented in the same way *i.e.* from the intermediate representations themselves. However for the ontology based on labels, the proposed hierarchy of elementary entities was manually validated prior to implementation and attributes at different levels of abstraction were arranged hierarchically rather than as siblings.

For each experiment the result was a multi-axial subsumption hierarchy, contained within and generated according to the respective underlying ontology, of source nursing interventions drawn from NIC.

2.3 Initial bench testing

Initial bench testing consisted of a comparison of the two generated hierarchies. The key criterion for comparison was whether and to what extent they captured hierarchical relationships

in the source terminology system. There are no explicit hierarchical relationships between nursing interventions within NIC. Therefore a test set of 73 implicit hierarchical relationships was identified manually. For example the NIC nursing intervention 'Bleeding Reduction: Antepartum Uterus' was considered to be a child of 'Bleeding Reduction'. Each of the two generated hierarchies was examined to see how many of the test set of implicit hierarchical relationships were captured.

2.4 Extending the ontology based on labels

The results of this initial bench testing (presented in Section 3) were used as the basis for rejection of the ontology based on informal definitions and for selection of the ontology based on labels for extension, further comparative bench testing and external evaluation. The ontology based on labels drawn from NIC was extended to include also nursing intervention components of two additional terminology systems, HHCC and OMAHA, using the same methodology as for the second experiment. In contrast to NIC these terminology systems are combinatorial in nature *e.g.* using OMAHA the notion of administering a medicine would be captured by combining the category 'Treatments and Procedures' with the target 'Medication administration'.

2.5 Comparative bench testing of the final ontology

Comparative bench testing took a randomly selected set of 30 terms drawn from the three source terminology systems, examined mappings within the UMLS Metathesaurus [18] to terms drawn from other systems, and compared these to mappings within the final extended ontology.

2.6 External evaluation of the final ontology

External evaluation provided an opportunity to assess the acceptability to 4 expert reviewers of the ontology and of its ability to mediate between diverse terminology systems. The generated hierarchy was used to identify:

- Hierarchical relationships within individual terminology systems *i.e.* from subsumee to nearest subsumer within the same terminology system
- Synonyms within individual terminology systems *i.e.* equivalent concepts within the same terminology system

- Mappings between terminology systems *i.e.* from subsumee to nearest subsumer in another terminological system
- Existing transformations between terminology systems *i.e.* equivalent concepts from different terminology systems
- Potential transformations between terminology systems *i.e.* close siblings from different terminology systems

For each set of relationships, pairs of related terms drawn from the generated hierarchy were presented in tabular form to the reviewers who were asked to indicate whether they deemed the relationships acceptable or not.

3. Results

3.1 Content and structure of the initial experimental ontologies

For the initial experimental ontology based on labels, the simplicity of the labels (in contrast to the relative complexity of informal definitions) was reflected in the resulting ontology; it consisted of 476 entities (compared to 992 in the ontology based on informal definitions) and 11 attributes (compared to 47). There were 131 compositional constraints (compared to 1766).

3.2 Results of initial bench testing

For the experimental ontology based on informal definitions, none of the hierarchical relationships within the test set were present within the generated hierarchy of nursing interventions (the total number of hierarchical relationships within the generated hierarchy was only 3). Analysis of the intermediate representations for the nursing interventions contained within the test set revealed several factors that prevented the formation of hierarchical relationships: a) structural differences between potential children and potential parents; b) different levels of specificity between potential parents and potential children; c) the absence of hierarchical relationships in the hierarchy of elementary GRAIL entities - unlike in the other cases this reflected on the development methodology rather than differences embodied within informal definitions.

Considerable work would be needed to overcome these factors *e.g.* a further round of normalisation to increase consistency, refinement of intermediate representations to loosen formal definitions for potential parents or to tighten formal definitions for potential children, and enforcement of appropriate classifications within the hierarchy of elementary GRAIL entities.

In contrast, for the initial experimental ontology based on labels there were 214 hierarchical relationships between nursing interventions. In many cases there was a richer structure within the generated hierarchy than within the test set, with several additional hierarchical relationships. As indicated previously the results of this initial bench testing were used as the basis for selection of the ontology based on labels for further development and evaluation.

3.3 Content and structure of the final extended ontology

The hierarchy, generated according to the final extended ontology, comprised all 1321 individual pre and post co-ordinated nursing interventions from all three source terminology systems. There were 2861 hierarchical relationships within the generated hierarchy. In terms of effort, there was a rapid tail-off when incorporating each of the two additional source terminology systems.

3.4 Results of comparative bench testing of the final ontology

Comparative bench testing demonstrated that in terms of *numbers* of mappings the final ontology proved to be more effective than the UMLS Metathesaurus - the final ontology suggested 45 mappings while the UMLS Metathesaurus suggested only 11. However, no assessment was made within this analysis of the *acceptability* of these mappings - the motivation behind external evaluation.

3.5 Results of external evaluation of the final ontology

The results of external evaluation, as summarised in Table 1, were mixed. For two of the reviewers involved in the external evaluation, the majority of hierarchical relationships within individual terminology systems derived from the final ontology were seen as acceptable. A third reviewer rejected the notion outright. In contrast, synonyms within individual terminology systems were rejected in the vast majority of cases by all three reviewers. Mappings and transformations between terminology systems were in the majority of cases seen as acceptable although one reviewer had significantly higher rates of rejection.

A major limitation in the interpretation of certain of the results was the lack of measure of agreement between different reviewers. While most of the analyses involved only one reviewer, the analysis of transformations did allow a limited but highly significant assessment, using Cohen's Kappa Coefficient, of inter-reviewer agreement. This showed that there was effectively no agreement between reviewers. While the broad range of analyses served to verify certain of the findings, they brought into question others by exposing inconsistencies on the part of individual reviewers and between reviewers.

4. Discussion

The use of informal definitions for nursing interventions resulted in highly complex intermediate representations. The hope for the ontology based on informal definitions was that the comparative richness of the informal definitions would be reflected within the ontology as a comprehensive set of hierarchical relationships; and that the rich generated hierarchy would facilitate the process of mapping between individual nursing interventions. This hope was dashed. In the first experiment the relatively rich representation embodied within the informal definitions required interpretation and allowed great freedom of expression. The need for systematic normalisation on initial intermediate representations demonstrated the difficulties associated with authoring consistently highly discursive definitional statements. The normalisation techniques used went some way in resolving unintentional differences between intermediate representations. However this largely manual process was arduous with little support provided by external resources.

The first experiment demonstrated that although it is indeed possible to derive an ontology from informal definitions for nursing interventions, the utility of such an ontology as a vehicle for mediation is highly questionable. To derive under this methodology an ontology with a richer hierarchy would require simpler sources with limited discursive content and a higher degree of consistency. The results of the second experiment provided evidence to support these claims. There was evidence that certain transformations had been carried out in the development of intermediate representations: the replacement of synonyms, the omission of certain constructs in the source labels, the expansion of relational adjectives and other 'packed' concepts and the use of 'role'-like constructs. However, there was no need, as in the first experiment, for harmonising

spelling and word form nor for transforming prepositions into links; the labels had already been normalised extensively during their development.

For the ontology based on labels the initial fear was that the simplicity and abstract nature of labels would result in fewer hierarchical relationships within the ontology. This fear was unfounded. The use of simpler sources had resulted in a more robust 'style guide' and greater coherence in the set of intermediate representations. The hierarchical organisation of attributes had further increased opportunities for subsumption and the manual validation of the entity hierarchy had contributed to the validity of subsumption. For these reasons, the initial ontology based on labels was subjected to a further round of development (*i.e.* the inclusion of nursing intervention components of HHCC and OMAHA), to further comparative bench testing and to external evaluation.

The final extended ontology covered all three source terminology systems in entirety *i.e.* every potential pre and post co-ordinated element was represented. The rapid tail off in terms of effort required to extend the ontology is perhaps indicative of a degree of commonality between the three terminology systems.

Nursing interventions were organised within the final ontology in a rich multi-axial hierarchy. The results of comparative bench testing demonstrated that the final extended ontology based on labels might overcome the limitations of other approaches to mediation. However, the results of more extensive external evaluation highlighted certain deficiencies in the development and evaluation methods employed.

The vast majority of mappings and transformations embodied within the final ontology were considered acceptable to reviewers. This is some indication of the usefulness of a logical ontology approach in exploiting similarities and resolving unmotivated differences between nursing intervention terminology systems. However a great many hierarchical relationships and synonyms, and several mappings and transformations were rejected on account of their informal definitions *i.e.* source labels were incomplete in terms of the meaning behind the nursing interventions they purported to represent. The heavy reliance on labels as the sole source for the final ontology resulted in errors. These errors could have been avoided to a large extent by paraphrasing labels (and validating the paraphrases) prior to more formal modelling.

Within the external evaluation, the rejection rate for synonymy within individual terminology systems (embodied within the final ontology) was high. This is particularly significant as certain sets of synonyms within the ontology had a profound impact on the acceptability of mappings and transformations between terminology systems. An analysis of the rejected mappings and transformations revealed that certain synonym sets appeared consistently to have caused problems. The approach taken within this study was useful in isolating these problem synonym sets.

Modelling activity within this study demonstrated that even relatively simple labels for nursing interventions can have complex hidden semantics, necessitating substantial nesting of entities and a relatively large number of attributes. Within this study the use of relatively simple intermediate representations that could be expanded automatically or semi-automatically into more complex GRAIL expressions greatly facilitated the manual modelling process.

The lack of agreement between reviewers provided some indication that there were significant ideological differences concerning both the nature of nursing interventions and the terminology systems that represent them. Such ideological differences certainly accounted for some of the difficulties in mapping. The range of analyses revealed a number of inconsistencies on the part of individual reviewers and between reviewers. In many cases these inconsistencies suggested that the ideological differences between the individual reviewers were not totally clear-cut; even firmly held beliefs could apparently change. One possible reason for this is simple reviewer error. However, many inconsistencies were not isolated cases; they occurred across individual analyses and between analyses. A second reason might be that many of the basic notions embodied within nursing terminology systems are inherently difficult to define - the study showed that comparatively vague notions like 'manage' are a major source of disagreement. A final possibility is that to accept similarities between terminology systems would in many respects be politically unacceptable. Each of the source terminology systems for this study has a large body of users and other stakeholders; the developers and users of each of the terminology systems have made large personal investments in their development and continued use; and the institutions built up around the terminology systems depend to a large extent on their individuality. Such political motivators require political solutions; they are certainly beyond the capability of a logical ontology.

One of the limitations of the approach to development and evaluation used within this study was that it did not seek to identify or promote consensus. This study has shown that the ontology is a suitable and useful vehicle for negotiation. A consensus type methodology *e.g.* some form of modified Delphi study, would facilitate evolution of the ontology and would obtain a more reliable consensus of opinion. Other researchers have reached similar conclusions [19]. This study has highlighted the deficiencies of a more summative approach.

As has already been suggested the source terminology systems for this study reflect the views of their respective developers; it is likely that they also represent in certain respects significant ideological differences across the nursing profession that are unlikely to be bridged by purely logical means. Indeed it might be wrong to do so: “Thinking only in computational terms, we run the risk of becoming focused exclusively on re-engineering all clinical work into formal behaviors that are suitable for computational treatment” [20]. If such ideological differences are indeed irreconcilable by any means, this raises real questions about the role of ‘classification’ in nursing in aggregating data from disparate sources [21]. However, there was no evidence within this study of such extreme differences. As such, a combination of a description logic-based approach and a consensus type development and evaluation methodology should over time lead to convergence.

5. Conclusion

In summary this study suggests that a logical ontology may be a useful tool in resolving a large proportion of differences between nursing intervention terminology systems; and that the use of nursing intervention labels as sources is highly productive, although some form of paraphrasing would be useful to capture missing semantics. However, the author believes that a complementary formative consensus type development and evaluation methodology, comprising discrete tasks, a range of reviewers, and a systematic assessment of agreement may improve the approach by helping to harmonise ideological differences.

Acknowledgements

This article represents a synthesis of work carried out as part of a Doctoral program conducted at the University of Manchester, UK under the supervision of Alan Rector. The author would like to thank Professor Rector for his guidance throughout the course of study. He would also like to

thank Susan Grobe, Derek Hoy, Stephen Kay, Katja Markert, Karen Martin, Joanne McCloskey Dochterman, Virginia Saba and Bonnie Webber for their support.

References

1. Spackman K, Campbell K, Cote R. SNOMED RT: A Reference Terminology for Health Care. In: Masys D, editor. AMIA 1997 Annual Symposium. Nashville: Hanley & Belfus, Inc.; 1997: 640-644.
2. Hardiker N, Rector A. Structural Validation of Nursing Terminologies. Journal of the American Medical Informatics Association 2001; 8(3): 212-221.
3. Chute C. Standards Move to Center Stage. MD Computing 1999; 16(1): 29-32.
4. Masarie F, Miller R, Bouhaddou O, Giuse N, Warner H. An Interlingua for Electronic Interchange of Medical Information: Using Frames to Map Between Clinical Vocabularies. Computers and Biomedical Research 1991; 24(4): 379-400.
5. McCormick K, Lang N, Zielstorff R, Milholland K, Saba V, Jacox A. Toward Standard Classification Schemes for Nursing Language: Recommendations of the American Nurses Association Steering Committee on Databases to Support Clinical Nursing Practice. Journal of the American Medical Informatics Association 1994; 1: 421-427.
6. Kim M, McFarland G, McLane A. Pocket Guide to Nursing Diagnoses. 5th ed. St. Louis: Mosby-Year Book Inc.; 1993.
7. McCloskey J, Bulechek G. Nursing Interventions Classification. 2nd ed. St. Louis: Mosby; 1996.
8. Martin K, Scheet N. The Omaha System: a pocket guide for Community Health Nursing. Philadelphia: W.B.Saunders Company; 1992.
9. Saba V. Home Health Care Classification of Nursing Diagnoses and Interventions. Washington D.C.: Georgetown University; 1994.
10. <http://www.ana.org/nidsec/nilang.htm>.
11. Ozbolt JG, Russo M, Stultz MP. Validity and Reliability of Standard Terms and Codes for Patient Care Data. In: Gardner RM, editor. 19th Annual Symposium on Computer Applications in Medical Care: Hanley & Belfus Inc.; 1995: 37-41.

12. Henry S, Holzemer W, Reilly C, Campbell K. Terms used by nurses to describe patient problems: can SNOMED III represent nursing concepts in the patient record? *Journal of the American Medical Informatics Association* 1994; 1: 61-74.
13. Rector A, Rossi-Mori A, Consorti F, Zanstra P. Practical development of re-useable terminologies: GALEN-IN-USE and the GALEN Organisation. *International Journal of Medical Informatics* 1998; 48: 71-84.
14. Rector A, Bechofer S, Goble C, Horrocks I, Nowlan W, Solomon W. The GRAIL concept modelling language for medical terminology. *Artificial Intelligence in Medicine* 1997; 9: 139-171.
15. Rogers J, Solomon W, Rector A, Pole P, Zanstra P, van der Haring E. Rubrics to Dissections to GRAIL. In: Pappas C, Maglaveras N, Scherrer J-R, editors. *Medical Informatics Europe (MIE 97)*. Thessaloniki, Greece: IOS Press; 1997: 241-5.
16. Rector A, Solomon W, Nowlan W, Rush T, Zanstra P, Claassen W. A Terminology Server for Medical Language and Medical Information Systems. *Methods of Information in Medicine* 1995; 34: 147-157.
17. Hardiker N, Webber B, Markert K, Rauch B. Supporting the development of formal models of nursing terminology. In: *Proceedings of Nursing Informatics*; Aukland; 2000.
18. Campbell K, Oliver D, Shortliffe E. The Unified Medical Language System: toward a collaborative approach for solving terminologic problems. *Journal of the American Medical Informatics Association* 1998; 5(1): 12-16.
19. Levy D, Dolin R, Mattison J, KA S, KE C. Computer-facilitated collaboration: experiences building SNOMED-RT. In: Chute C, editor. *A Paradigm Shift in Health Care Information Systems: Clinical Infrastructures for the 21st Century*; 1998: 870-874.
20. Coiera E. When Conversation Is Better Than Computation. *Journal of the American Medical Informatics Association* 2000; 7(3): 277-286.
21. Zielstorff RD, Hudgings CI, Grobe SJ. *Next-Generation Nursing Information Systems: Essential Characteristics for Professional Practice*. Washington, DC: American Nurses Publishing; 1993.

Table 1: Summary of results from external evaluation of the final ontology

	Hierarchical relationships accepted	Synonym pairs accepted	Mappings accepted	Transformations accepted	Potential transformations accepted
Reviewer #1	88%	17%	-	86%	97%
Reviewer #2	72%	0%	17%	56%	40%
Reviewer #3	0%	0%	97%	71%	65%
Reviewer #4	-	-	89%	-	-