Bridging the gap between narrative text guideline and its computer-interpretable version: taking into consideration and supporting the user perspective

Patrick Martini

Institute of Software Technology & Interactive Systems
Vienna University of Technology, Vienna, Austria
martini@ifs.tuwien.ac.at
http://ieg.ifs.tuwien.ac.at

Abstract. This proposal focuses on the research question on how to develop a system that produces a formalized guideline. The output of this research project on the one hand, helps the medical personal in consulting and in their decision-making process and, on the other hand, fills the gap between the CPGs and their computer-interpretable representations. We use a document-centric model approach, which goal is to lose as minimum information as possible, preserving so the original "guideline spirit". We will create a formalization that can be used as a connecting bridge between computer-interpretable guideline representation and the narrative text of a guideline. From our perspective a guideline should have two different formalizations: one that is used and created for end users (e.g., consulting guideline for physicians) and another one that can more easily be executed and interpreted by a computer. Hence, we consider the problem from another perspective as regards to the current approaches.

1 Introduction

The evidence-based medicine/healthcare (EBM) [26] movement has been gaining ground quickly over the past few years promoted by politicians, clinicians, and management in order to improve the quality and consistency in healthcare reducing costs where possible. EBM also influenced the development of clinical practice guidelines (CPGs) – also known as medical guidelines – to arrive at higher-quality of CPGs compared to the consensus-based development.

The standard definition of CPGs is that of Field and Lohr in 1990 [8] stating that they are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances". Such guidelines are designed to assist health care practitioners in the prevention, diagnosis, treatment, and management of specific clinical conditions.

CPGs, used correctly, have been proven to be capable of improving patient care [15] and supporting health care practitioners in their decision-making process. As yet they have neither a huge diffusion nor do they have a big impact on medical practice, but their importance is growing constantly. There are different reasons for these poor results; neither the format of narrative text of the current CPGs nor the natural distrust of human beings in new things does help the diffusion.

CPGs are often characterized by their limited use in practice: rarely, practice behavior is changed by them [34]. As analyses have shown passive dissemination of guidelines (e.g., publication in a medical journal) has not been successful in changing behavior. However, when additional active implementation strategies have been used then one can observe changes in practice patterns [7].

The new focus of EBM is directed towards the development of effective strategies for translating evidence into practice. For supporting this process we have to develop tools that help the clinicians in using CPGs. Using a marketing terminology we could say that our aim is to render the guidelines winning.

This is a process that involves a lot of disciplines and our aim is to cover the computer-science part. To address the latter problematic, several groups have created computer-interpretable guideline representations (an exhaustive overview of different approaches can be found in [23]).

From our perspective there exists a gap between the narrative versions of guidelines and their computer-interpretable counterparts, in the sense that they cover only the executing part forgetting all the other elements that are useful for physicians for their final decision, such as strength of recommendation, level of evidence, etc.

In my PhD-thesis I am therefore going to study and analyse the following research question:

How can we develop a system that given a guideline produces a formalized guideline that, on the one hand, helps the medical personal in consulting and in their decision-making process and, on the other hand, fills the gap between the CPGs and their computer-interpretable representations?

In other words, we study the problem by focusing mainly on the physicians' perspective in how to utilize successfully the guidelines in their daily activities.

In the next section we give an overview of the current related approaches. In Section 3 we present the preliminary terminology we have adopted with a particular focus on our personal concept of guideline formalization. The description of our approach can be found in Section 4 followed by work performed so far, future working steps and conclusions.

2 Related Work

Currently, two kinds of approaches exist to model computer-interpretable CPGs: (1) model-centric techniques and (2) document-centric techniques. In the former, a conceptual model is formulated by domain-experts. Thus, the relationship between the model and the original document is only indirect. Representatives of this techniques are *AsbruView* [19], *Protégé* [10], *Arezzo*, and *Tallis* [30]. In document-centric approaches markup-based tools are used to systematically mark-up the original guideline in order to generate a semi-formal model of the marked text part. Representative tools are the *GEM Cutter* [24], *Stepper* [25], the *Document Exploration and Linking Tool / Addons (DELT/A)* [32], *Uruz*, part of the *Digital electronic Guideline Library (Degel)* framework [29] and *Clinical Practice Guideline Reference Architecture (CPG-RA)* the revisited version of the previous CPGA [5].

MHB (Many Headed Bridge between CGP formats) [28] bridges the gap between the text-version of a CGP and its formal counterpart.

All these approaches accomplish their task in a manual manner. We try to disburden the involved users by automating parts of this task and improving their communication by providing information needed by them. We will explain our approach in Section 4.

3 Preliminary Considerations

First and foremost, to us it seems appropriate to define some terminology and concepts on which all our work is based. Usually, a medical guideline comes in narrative text in which we can recognize a logic between different elements. In particular, for us, a guideline is a set of recommendations which, in turn, is formed by a number of chunk texts. A chunk text is the minimal component of the original text in which we can subdivide the guideline following particular criteria. The relation between recommendation and chunk text is crucial. It is vital to know from which recommendation a chunk text comes from. We define an entity as a set of chunk texts with a correlation between them.

We can imagine a guideline like a directed graph in which edges (relation) connect to nodes (recommendation) in only one direction.

We consider only edges between recommendations feasible. Is it also possible to have relations between chunk texts/entities of different recommendations ?

In our work the formalization of the CPGs has a central role so we have to think what formalizing a CPGs means. A common answer could be: Formalizing a medical guideline means to process the original document by using one of the guideline representation languages in order to transform it in a format that can be easily processed by a computer. With this approach, the process is bound to the destination language.

Our point of view is slightly different. For our formalizing process we decide to wipe out from our mind the characteristic of a particular destination language (e.g., Asbru) and focus instead on the health care practitioners' point of view trying to understand the support they really need putting aside the destination languages and their possible limitations.

Before we can proceed, fundamental questions require an answer which form one of the major issues of our research project. Their outcome will lead the next steps. More precisely, these questions are:

Which are the useful elements at stake in the medical decision-making process?
How can we classify and represent these elements bearing in mind the requirements for medical modelling languages expressed by Tu and Musen [31] in order to make the shift to a computer-interpretable guideline representation as smoothly as possible?

4 Our Approach

Our approach is influenced by the tenet that a medical guideline decides nothing and it is only an instrument to facilitate the decision-making process that physicians have to face every day. Daniel J. Friedland states [9]: "Optimal decision-making requires that physicians identify each possible strategy, accurately predict the probability of future events, and balance the risks and benefits of each possible action, all while tailoring to the values of individual." So the guideline information is going to complete the personal knowledge of a physician. For this reason, if we do not consider the part of guideline that contains an action (hence, being the most important part of a guideline), we cannot standardize which information is useful and which is not, as it depends on the user who is consulting the guideline. Consequently, we have adopted a document-centric model [5] approach, which goal is to lose as minimum information as possible, preserving so the original "guideline spirit". This kind of approaches permit us in the future to extend the CPGs with the personal knowledge-annotations of healthcare practitioners. Hence, we do not decide which information is important, but rather the end user can adapt it to his/her pleasure.

We have seen that the approaches described in Section 2 accomplish their task with a manual modelling which is error prone and hugely time-consuming. Thus, we use a semi-automatic approach which reduces the manual part as much as possible. We set our problem in the area of Natural Language Processing (NLP). In order to address the problem, we build different components that operate in a pipeline fashion (see Figure 1). This pipeline takes a guideline as input and the transformed version as output. This pipeline should be flexible and adaptable permitting to interchange each element. This can be achieved by means of an intermediate XML-representation of original guidelines. In each step the guideline is augmented by elements and is transformed in ever more formal version.

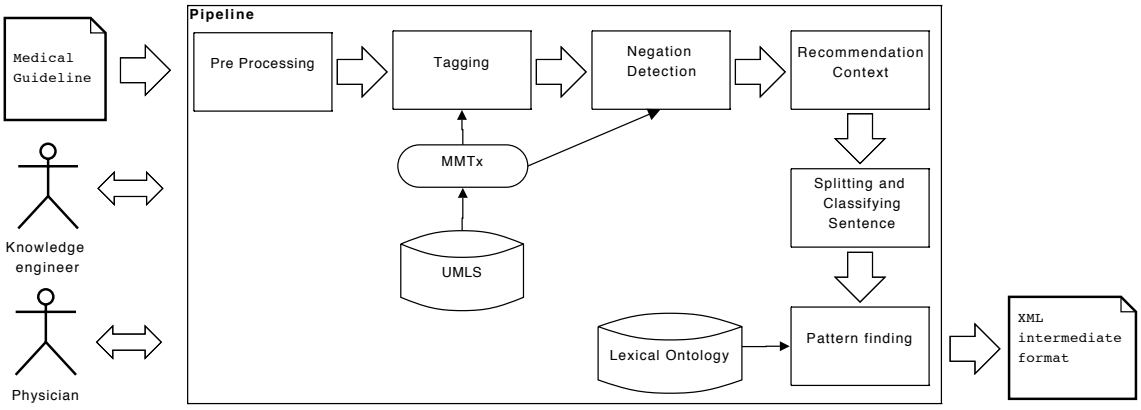

Fig. 1. Pipeline Overview

The active actors involved in this process are the following ones: knowledge engineers and healthcare practitioners. Because of their different backgrounds and aims we can simplify their roles by saying that the healthcare practitioner judges, confirms, and eventually rejects the guideline formalization produced by the knowledge engineer. In each step of the formalization process we try to answer the following question:

What are good presentations and visualizations of the extracted information in order to help these actors in formalizing, consulting, and interacting with a medical guideline ?

Answering this question gives us useful elements for our formalizing process as well as for future work regarding the development of a framework that can be a valid support during all the guideline formalization phases and that in a last instance makes the consultation and the visualization a rewarding experience.

For our scope, we selected 20 treatment/management cancer guidelines from National Guideline Clearinghouse^{TM}(NGC[1]), a public resource for evidence-based and consensus-based clinical practice guidelines. The intrinsic complexity of the branch of treatment/management cancer gives us a good "training" set for our work.

Briefly, we now describe the steps of our pipeline system.

[1] http://www.guideline.gov/

4.1 Pre-processing

Exploiting the HTML-format of the guidelines, we obtain a coarse first XML-representation of the guideline that constitute the basis for the following steps. In particular, we identify each recommendation and extracted for each one the information regarding the strength of recommendation, the level of evidence, the benefits (Grade approach [2]) and the title of the recommendation (in case this information is available). Seeing that each organization, who develops CPGs, uses a particular HTML-template we implemented a Python web-scrapper for each.

4.2 Medical Concepts and Semantic Types Tagging

We argue that a reliable and well-considered stage of tokenisation is a necessary and highly valuable basis for a higher-level NLP process. In this step, we identify medical concepts present in a single recommendation [22]. For this purpose, we use the Unified Medical Language System (UMLS) combining it with the MetaMap Transfer (MMTx) program [1]. This is an important step for the following part.

4.3 Negation Detection

A negation detection system is a crucial and delicate part for a formalization of a medical guideline and for each system that deals with Information Extraction on medical content in general. Without it one could misunderstand the extracted information therein enclosed and this could influence negatively the final results. Negation is a fundamental operation that can invert the sense of a sentence, a condition etc. While negation in predicate logic is well defined and syntactically simple (for example "not a" is simply the polar opposite of "a"), negation in natural language is complex and has been philosophized about for hundreds of years [13]. This has been discussed in detail by Horn and Kato [14]. The following examples give an indication how complex the problem could be [21]:

1. I'm not tired
2. I'm not a bit tired (=I am not at all tired)
3. I'm not a little tired (=I am quite tired)

It emerged that the negation with "not", the negation of passive sentences and the double negation create some problems [17]. Therefore, we have to analyse the text in order to define pattern-based rules. Fortunately, the negation in medical narrative is more direct and straightforward and rarely, it crosses the sentence bound. We evaluate NegEx [4] and Netfinder [21] algorithms.

4.4 Recommendation Context

Can we deduce from the latter two steps the context of an entire guideline? And more precisely, can we also deduce a context for each recommendation? If not (or in cases of partial contexts), can we use the general context for deducing the context of the recommendation (e.g., population)?

Can we create a summary of a guideline by merging the various contexts?
A summary is like a guideline-photo. By giving a glance on it, one can have a general idea about what is going on in the guidelines. For example: population, number of recommendations, the type of guideline, recommended drugs, and so on. A summary is useful not only in the final stage (consultation guideline) but gives also hints for the coming steps. However, this task need to developed further. At the moment, we have only a vague idea.

4.5 Splitting and Classifying the Sentence

This step and the following ones aim at facilitating an eventual formalization in Asbru. First of all, we assembly the chunk text/entity in sentences: so now a recommendation is a set of sentences. We try to classify roughly the sentence in "action" and the "auxiliary information". We consider the term action in its broadest definition. The auxiliary information obtains a level of visibility that allows information to be rendered personalized in the consultation phase according to the physicians' preferences.

4.6 Pattern finding

By combining lexical ontology [27] and the UMLS concept we detect, we try to extract linguistic patterns and consequently relations between different actions composing a recommendation. The action now is pured from all parts which are not necessary, like stop words, etc. In this step, we shift from an action in narrative text to a formalized action in which the semantic type and the concepts of the original action is embedded.

This is useful for two reasons. First, it helps the migration of our XML-guideline to a computer-interpretable guideline. Second, it could be used to grasp the difference between two versions of the same action (e.g., "living guidelines") replacing the text matching with a semantic one.

5 Work Performed So Far

The first step of my thesis work was to understand what evidence-based medicine and connected to it CPGs are about, and the philosophy that stands behind [26, 34, 33]. Afterwards, we investigated the State-of-The-Art in guideline formalization methods and tools (see Section 2). In addition, we studied the State-of-The-Art in Information Extraction [12, 3, 11] and also we have a look at Natural Language processing [18, 6] for less structured information (e.g., medical narrative text) – in particular applied to the medical area [20].

As pointed out earlier in the proposal, our main concern is the physicians' perspective. Therefore, we went to the G-I-N[2] Conference 2006 in Budapest in order to meet with physicians that are interested in evidence-based medicine. It emerged that every physician has its own ideas on which parts of a guideline are important and how to use it. Hence, we decided to make a formalization which needs to be very flexible in order to satisfy the preferences of every single physician.

6 Future working steps

The next steps that need to be taken in the following months are:

- Defining a robust XML format and the relative Xschema that at the same time should be robust and elastic for future changes;
- Defining a system to identify univocally the different elements composing the guideline. Doing so will allow us to gather the changes of elements during the formalization process;
- Tagging medical concepts and semantic types;
- Detecting negation;
- Defining recommendation context;

[2] http://www.g-i-n.net/

– Splitting and classifying the sentence;
– Finding patterns.

Next to these steps, we also need to work on evaluation. The evaluation part is split in two phases. The first phase forms the simple part and consists in evaluating for each step its precision and recall scores. For this, we need at least one knowledge engineer and one physician. The second phase evaluates the system's logic and semantic. In this respect, we foresee to build a voting system in which three physicians will be involved. This works as follows: two physicians vote for the correctness of each piece of the transformed recommendation, while the third one can intervene when the others are not of the same opinion.

Another aspect we need to evaluate is whether our representation format is to be compatible with existing guideline modelling methods (GMMs), meaning that the information can be transformed into such languages as Asbru, Proforma, and so on. To prove the compatibility, we will show that transformation from our intermediate format into Asbru format is possible. During this process we can also recognize possible gaps presented in these languages that, for us, that need to be plugged (e.g., in Asbru the information about the evidence can only be represented in a text label without a semantic).

7 Conclusions

We will create a formalization that can be used as a connecting bridge between computer-interpretable guideline representation and the narrative text of a guideline. From our perspective a guideline should have two different formalizations: one that is used and created for end users (e.g., consulting guideline for physicians) and another one that can more easily be executed and interpreted by a computer. Hence, we consider the problem from another perspective as regards to the current approaches.

First, we look at the users and their necessities – in particular we focus on physicians – and only afterward we regard the technical aspect. This kind of proceeding gives us a broader overview of the problem(s) we are facing.

Hence, this project tries to satisfy the wishes of physicians and consequently, the CPG dissemination might be improved. At this particular moment, we are in the preliminary stages of our project.

Acknowledgments This work is supported by "Fonds zur Förderung der wissenschaftlichen Forschung FWF" (Austrian Science Fund), grant L290-N04.

References

1. A. R. Aronson. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In *Proc AMIA Symp*, pages 17–21, 2001.
2. David Atkins, D. Best, Peter A. Briss, Martin Eccles, Y. Falck-Ytter, Signe Flottorp, Gordon H. Guyatt, Robin T. Harbour, M. C. Haugh, D. Henry, Suzanne Hill, Roman Jaeschke, G. Leng, Alessandro Liberati, Nicola Magrini, James Mason, P. Middleton, J. Mrukowicz, Dianne O'Connell, Andrew D. Oxman, Bob Phillips, Holger Schünemann, Tessa Tan-Torres Edejer, H. Varonen, Gunn E. Vist, John W. Williams Jr., S. Zaza, and The GRADE Working Group. Grading quality of evidence and strength of recommendations. *British Medical Journal*, 328(7454):1490–1498, June 2004.

3. Mary Elaine Califf and Raymond J. Mooney. Relational learning of pattern-match rules for Information Extraction. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 328–334, Orlando, FL, July 1999.
4. Wendy W. Chapman, Will Bridewell, Paul Hanbury, Gregory F. Cooper, and Bruce G. Buchanan. A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of Biomedical Informatics*, 5:301–310, 2001.
5. Stephen Chu. Guideline representation formalism and electronic decision support systems: Addressing the guideline – implementation gap. *Health Care and Informatics Review Online*, March 2005.
6. H. Cunningham, K. Humphreys, R. Gaizauskas, and Y. Wilks. Software infrastructure for natural language processing, 1997.
7. D. A. Davis, M. A. Thomson, A. D. Oxman, and R. B. Haynes. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials. *JAMA*, 268(9):1111–1117, 1992.
8. Marilyn J. Field and Kathleen N. Lohr, editors. *Clinical Practice Guidelines: Directions for a New Program.* National Academies Press, Institute of Medicine, Washington DC, 1990.
9. Daniel J. Friedland. *Evidence-Based Medicine: A Framework for Clinical Practice*. McGraw-Hill Professional, 1998.
10. John H. Gennari, Mark A. Musen, Ray W. Fergerson, William E. Grosso, Monica Crubézy, Henrik Eriksson, Natalya F. Noy, and Samson W. Tu. The Evolution of Protégé: An Environment for Knowledge-based Systems Development, volume = 58, year = 2003. *International Journal of Human Computer Studies*, (1):89–123, 2003.
11. William R. Hersh. *Information Retrieval. A Health and Biomedical Perspective*. Springer, second edition, 2003.
12. Richard D. Holowczak and Nabil R. Adam. Information Extraction based Multiple-Category Document Classification for the Global Legal Network. In *Proc. of the 9th Conference on Innovative Applications of Artificial Intelligence*, pages 992–999. AAAI Press/MIT Press, 1997.
13. Laurence R. Horn. *A Natural History of Negation*. University of Chicago Press, Chicago, 1989. Reissued 2001 by CSLI.
14. Laurence R Horn and Yasuhiko Kato. *Negation and Polarity: Syntactic and Semantic Perspectives*. Oxford University Press, Oxford, 1999.
15. Mary E. Johnston, Karl B. Langton, R. Brian Haynes, and Alix Mathieu. Effects of computer-based clinical decision support systems on clinician performance and patient outcome: a critical appraisal of research. *Ann Intern Med*, 120(2):135–142, 1994.
16. Katharina Kaiser. *LASSIE – Modeling Treatment Processes Using Information Extraction*. PhD thesis, Institute of Software Technology & Interactive Systems, Vienna University of Technology, 2005.
17. Jung-Jae Kim and Jong C. Park. Extracting contrastive information from negation patterns in biomedical literature. *ACM Transactions on Asian Language Information Processing (TALIP)*, 5(1):44–60, 2006.
18. Margaret King. Evaluating natural language processing systems. *Communications of the ACM*, 39(1):73–79, January 1996.
19. Robert Kosara and Silvia Miksch. Metaphors of Movement: A Visualization and User Interface for Time-Oriented, Skeletal Plans. *Artificial Intelligence in Medicine, Special Issue: Information Visualization in Medicine*, 22(2):111–131, May 2001.
20. S. M. Meystre and P. J. Haug. Comparing natural language processing tools to extract medical problems from narrative text. *AMIA Annu Symp Proc*, pages 525–529, 2005.
21. Pradeep G. Mutalik, Aniruddha Deshpande, and Prakash M. Nadkarni. Use of general-purpose negation detection to augment concept indexing of medical documents a quantitative study using the umls. *American Medical Informatics Association*, 8(6):598–609, 2001.
22. P. Nadkarni, R. Chen, and C. Brandt. Umls concept indexing for production databases: a feasibility study. *J Am Med Inform Assoc*, 8(1):80–91, 2001.
23. Mor Peleg, Samson W. Tu, Jonathan Bury, Paolo Ciccarese, John Fox, Robert A. Greenes, Richard Hall, Peter D. Johnson, Neill Jones, Anand Kumar, Silvia Miksch, Silvana Quaglini, Andreas Seyfang, Edward H. Shortliffe, and Mario Stefanelli. Comparing Computer-Interpretable Guideline Models: A

Case-Study Approach. *Journal of the American Medical Informatics Association (JAMIA)*, 10(1):52–68, Jan-Feb 2003.
24. Kristi-Anne Polvani, Abha Agrawal, Bryant Karras, Aniruddha Deshpande, and Richard Shiffman. *GEM Cutter Manual*. Yale Center for Medical Informatics, 2000.
25. Marek Růžička and Vojtech Svátek. Mark-up Based Analysis of Narrative Guidelines with the Stepper Tool. In Katharina Kaiser, Silvia Miksch, and Samson W. Tu, editors, *Computer-based Support for Clinical Guidelines and Protocols. Proceedings of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101 of *Studies in Health Technology and Informatics*, pages 132–136, Prague, Czech Republic, 2004. IOS Press.
26. David L Sackett, William M C Rosenberg, J A Muir Gray, R Brian Haynes, and W Scott Richardson. Evidence based medicine: what it is and what it isn't. *BMJ*, 312(7023):71–72, 1996.
27. Radu Serban, Annette ten Teije, Frank van Harmelen, Mar Marcos, and Cristina Polo-Conde. Extraction and use of linguistic patterns for modelling medical guidelines. *Artif. Intell. Med.*, 39(2):137–149, 2007.
28. Andreas Seyfang, Silvia Miksch, Mar Marcos, Jolanda Wittenberg, Cristina Polo-Conde, and Kitty Rosenbrand. Bridging the gap between informal and formal guideline representations. In Gerhard Brewka, Silvia Coradeschi, Anna Perini, and Paolo Traverso, editors, *European Conference on Artificial Intelligence (ECAI-2006)*, volume 141, pages 447–451, Riva del Garda, Italy, 2006. IOS Press.
29. Yuval Shahar, Ohad Young, Erez Shalom, Alon Mayaffit, Robert Moskovitch, Alon Hessing, and Maya Galperin. DEGEL: A hybrid, multiple-ontology framework for specification and retrieval of clinical guidelines. In Michel Dojat, Elpida Keravnou, and Pedro Barahona, editors, *Proceedings of the 9th Conference on Artificial Intelligence in Medicine in Europe, AIME 2003*, volume 2780 of *LNAI*, pages 122–131, Protaras, Cyprus, 2003. Springer Verlag.
30. R. Steele and John Fox. Tallis PROforma Primer – Introduction to PROforma Language and Software with Worked Examples. Technical report, Advanced Computation Laboratory, Cancer Research, London, UK, 2002.
31. Samson W. Tu and Mark A. Musen. A flexible approach to guideline modeling. In *Proceedings of the AMIA Annual Symposium*, pages 420–424, Washington DC, 1999.
32. Peter Votruba, Silvia Miksch, and Robert Kosara. Facilitating knowledge maintenance of clinical guidelines and protocols. In Marius Fieschi, Enrico Coiera, and Yu-Chuan Jack Li, editors, *Proceedings from the Medinfo 2004 World Congress on Medical Informatics*, pages 57–61. IOS Press, 2004.
33. Brandi White. Making evidence-based medicine doable in everyday practice. *Family Practice Management*, 11(2), February 2004.
34. Steven H Woolf. Evidence-based medicine and practice guidelines: an overview. In *Cancer Control 7*, 4, pages 362–367, 2000.