ELSEVIER Artificial Intelligence in Medicine 12 (1998) 153–168

Artificial Intelligence in Medicine

Distributed cognition and knowledge-based controlled medical terminologies

J.J. Cimino *

Department of Medical Informatics, Columbia University College of Physicians and Surgeons, New York 10032, USA

Received 31 October 1996; received in revised form 31 August 1997; accepted 1 September 1997

Abstract

Controlled medical terminologies (CMTs) are playing central roles in clinical information systems and medical knowledge resource applications. As these terminologies grow, they are able to support more complex tasks but require more intensive efforts to create and maintain them. Several terminologies are evolving into knowledge bases of medical concepts. The knowledge they include is being used to support distributed cognition in two forms: complex medical decisions involving multiple people and applications, and coordination of maintenance of the terminologies themselves. © 1998 Elsevier Science B.V.

Keywords: Controlled medical terminology; Distributed cognition; Vocabulary maintenance; Clinical system integration; Clinical decision support

1. Introduction

No health care practitioner today is expected to be capable of carrying out the full spectrum of intellectual processes required for modern medicine. Not only is the body of medical knowledge beyond the capacity of an individual mind; today's standard of care requires us to call on others whose expertise and experience differ

* Present address. Columbia-Presbyterian Medical Center, Atchley Pavilion, Room 1310, New York 10032, USA. Tel.: +1 212 3058127; fax: +1 212 3053302; e-mail: James.Cimino@columbia.edu

PII S0933-3657(97)00048-1

from our own if they promise an advantage to the patient. This expertise is typically found in consultants and specialists, but it may also reside in information resources and expert systems. We can define distributed cognition to occur when intellectual processes are shared among multiple participants, especially in order to solve a particular task in a particular context. Efficient distributed cognition occurs when each participant is relieved of some part of the process, with a minimum of redundancy and with a minimum of effort to coordinate the participants. Efficiency, in turn, depends upon the ability to transfer information, especially contextual information, effectively (accurately and efficiently) among the participants. Specifically, information about the patient and his problems must be transferred from the primary practitioner to the consultant (human or computer) and relevant information (such as diagnostic and therapeutic recommendations) must be transferred back—all with a minimum of explanation and inaccuracy. When computer systems are part of this process, we are almost always required to involve controlled medical terminologies (CMTs)[1].

CMTs, in one form or another, are at the heart of most medical systems. In many cases, these terminologies are little more than word lists, used for capturing information about patients or indexing medical knowledge. As medical systems have grown in complexity and sophistication, more demands are being placed on the terminologies used. For example, an application used for keeping track of patient problems might include a list of terms for selection by a physician in order to simplify the data input process. However, if the list contains redundant or ambiguous terms, the data captured with these terms become unreliable and of diminished value. Weaknesses in CMTs become apparent with use, and the emergence of Internet-based medical systems is putting CMTs to the test, as systems are integrated and patient information is transferred among them. Such integration demands that CMTs be precise and well-disciplined, for they are of little use if the receiving system cannot understand what the sending system is saying.

CMTs are evolving to meet these demands. One way in which they are changing is that where previously they had been used for such tasks as representing the concepts in an expert system's knowledge base, they are now becoming knowledge bases themselves, containing definitional information [7,12,47]. The knowledge in CMTs serves a variety of purposes: as a way for humans to understand the explicit meanings of concepts in the CMT, to support inferencing by expert systems, and to aid in the maintenance of the CMT itself. This evolution, however, is imposing new cognitive tasks related to the creation and maintenance of CMTs-tasks which are complex enough and tedious enough that they too may benefit from a distributed approach. The intent is that advanced, knowledge-based CMT development, distributed or otherwise, will pay off with more efficient distributed cognition in health care. This paper reviews some experience to date with such efforts.

[1] Several formal terminological terms are used throughout this paper. Controlled terminology refers to a collection of allowable names, called terms. If the terminology is concept-based, then the terms correspond to particular meanings. When multiple names map to the same concept, they are referred to as synonyms, with one name usually chosen as the preferred term for the concept. Information which describes the concept (or meaning) is called the definition.

When CMTs become knowledge bases, they become directly relevant to the topic of cognition. This paper will address two aspects of these cognitive capabilities as they relate to the theme of distributed cognition: the distributed development of the CMT-knowledge base, and the use of knowledge-based CMTs to support distributed cognition in medical care.

2. Distributed development knowledge-based CMTs

2.1. Why are CMTs evolving into knowledge bases?

Traditionally, developers of medical applications created their own CMTs on an as-needed basis. In general, little attention was paid to this aspect of the applications and the CMTs were little more than lists of terms with which users could fill in fields of a form or respond to questions. As applications became more complex and began to cover broader medical domains, application developers suddenly found that they needed larger CMTs. For example, expert systems which collect patient signs and symptoms to perform differential diagnosis require term lists numbering in the thousands [2,35]. As large as these lists have become, they are much too small for tasks where greater expressivity was needed, such as electronic medical record keeping. Application developers naturally seek to simplify their task by attempting to adopt CMTs developed by others. However, the CMTs created for a specific application were generally found to be unusable. For example, a CMT created for an electronic medical record was found unsuitable for use in an diagnostic expert system despite the fact that both systems were created in the same laboratory [51].

The need for reusable CMTs led to the creation of large, application-independent terminologies which, it was hoped, would be usable in many settings. Among these were the US National Library of Medicine's (NLM's) medical subject headings (MeSH) [38], the College of American Pathologists' systematized nomenclature of medicine (SNOMED) [18], and the Gabrieli nomenclature [22]. Although not created specifically for use by computer systems, developers also attempted to use the International Classification of Diseases, 9th edition, with Clinical Modifications (ICD9-CM—a US extension of ICD9) [49]. In the United Kingdom, the Read Codes were made available for use in record keeping systems and eventually mandated for use by the National Health Service. In the Netherlands, the Elias system [50] was developed for use in doctors' office systems and adapted the International Classification of Primary Care [29] for this purpose. The NLM has been developing the Unified Medical Language System to bring many of these CMTs together into a single resource [30].

Despite the availability of large CMTs and the clear desire not to reinvent the wheel, application developers have been slow to adopt these CMTs for their own use. There are many reasons for this resistance; some of the serious ones stem from the fact that the meanings of the terms in the CMTs are not made explicit. As a result, these meanings are left open to interpretation by potential users. Close examination of the content of publicly available CMTs shows that they are plagued with redundancy, ambiguity, and vagueness [12]. Medical informatics researchers have hypothesized that these problems can be overcome through the inclusion of definitions in the CMT. Furthermore, these definitions should be in a computer-manipulable form so that smart CMT tools can be created which can help users understand the content to locate appropriate terms for their intended meaning [7,12].

These hypotheses have lead CMT developers to attempt to provide definitions through the use of structured, named interrelationships among concepts. So, for example, disease terms would have specific relationships to other terms in the CMT, indicating the causative agent (etiology) and the body location involved (site). The result can be viewed as a set of frames, with named slots filled with values that refer to other frames, or it can be viewed as a semantic network where terms are nodes and links are named relationships among the terms [3]. A variety of notation systems are used, with conceptual graphs being one of the most common [4,6]. The SNOMED editorial board has recently committed to the inclusion of such definitional knowledge in its next major release [48]. In a separate effort, a group of independent system developers and users has created the logical observations, identifiers, names and codes (LOINC) which is based on the construction of terms using a strict definitional structure [21].

2.2. Using the knowledge in a CMT for maintenance

The definitional knowledge in CMTs is perhaps best put to the test by tasks involving the maintenance of the vocabulary itself. The editing facilities envisioned for the VOSER project, for example, will make use of vocabulary dependencies to constrain vocabulary maintainers in order that they not disrupt internal consistency [47].

A typical task is the addition of new terms to the CMT. Definitional information can be used to help determine if the CMT already contains a synonymous term to which the new term can be added or, if not, can propose where and how the new term should be added. A second task is the proper subsumption of terms by other terms where 'is-a' or subclass relationships should exist. For example, if the CMT contains the terms 'infectious disease' and 'bacterial pneumonia', it would be important to recognize that the latter is subsumed by the former. If the term 'lung disease' is later added to the CMT, a second 'is-a' link should be added between it and terms already in other classes, such as bacterial pneumonia. Manual performance of these tasks is tedious and may not be reliable in practice. Finding ways for intelligent tools to help with the task is a form distributed cognition.

One set of experiments with this type of maintenance capability has been conducted at Columbia University, with the development of the Medical Entities Dictionary (MED) [12], a CMT used for coding clinical data collected from ancillary systems and stored in the central data repository of the Presbyterian Hospital [28]. A simple example of the MED structure can be seen in Fig. 1, showing the definition of the term 'fasting glucose test' through its links to other terms. The first use of this knowledge was to allow for the automated classification of laboratory terms, including the discovery of natural classes among the terms [13]. In that effort, 526 test terms were organized into a structure of 36 classes, based on knowledge of the specimens and analytes (the substances measured by the tests).

For several years, this knowledge was used to support ongoing maintenance of the CMT, aiding in the addition of, among other things, 224 new test terms. Then, in 1994, a new laboratory system was installed at Presbyterian Hospital. This new system prompted the laboratory personnel to develop an entirely new terminology. On the surface, the old and new terminologies appeared incompatible. However, through the knowledge modeling process, the 840 new test terms were successfully integrated into the MED in the 1 month period between vocabulary creation and system completion [15]. Success, in this case, was defined by the fact that when new laboratory data started being received by the central repository, they were stored, retrieved, displayed, and used for automated decision support without interruption in service. Thus, the editing tools took advantage of the knowledge in the MED to

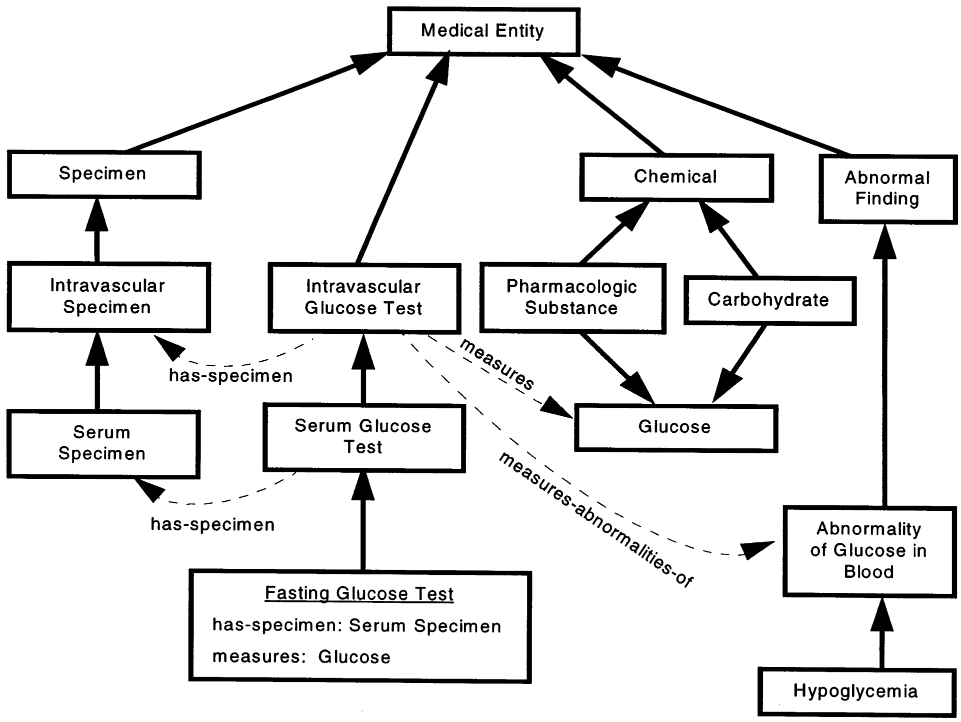

Fig. 1. Fig. 1 shows a simplified view of relationships among concepts in the MED semantic network. Solid arrows are 'is-a' links, broken arrows are nonhierarchical semantic links. The definition for one concept 'fasting serum glucose' is shown as a frame. It's location in the MED hierarchy is determined by its relationships to the concepts serum specimen and glucose. Not shown is the relationship between this concept and abnormalities of glucose in blood, which it inherits. Note that multiple hierarchies, such as that shown for glucose, are allowed.

perform some inferencing (and hence cognitive) tasks, off-loading the human editor efficiently and effectively.

Meanwhile, European researchers were facing a much greater terminology integration problem. In order to share clinical information among the members of the Economic Community, CMT developers needed to merge terminologies not only of different systems but different languages. The Galen project arose to take on this task through the use of a very explicit compositional grammar which attempts to represent everything that is 'sensible to say' [44]. Using the knowledge contained in GALEN, developers are able to test for four functional types of correctness and completeness: conceptual, linguistic, inferential, and pragmatic [45].

In the United Kingdom, the Read Thesaurus is being expanded to include semantic definitions. While these definitions are proving useful for supporting evaluation tasks [42], the development of complete definitions many not be practical for a large percentage of terms [5].

One research group at IBM has taken a formal knowledge representation system called KRep and adapted it for use in maintaining knowledge-based CMTs [33]. In this approach, not only is definitional information possible, it is required in order to convert 'primitive terms' into 'defined terms'. Once a term is defined, the system can automatically determine if its definition matches that of any other term, pointing to possible redundancies or places where additional 'differentia' are needed to distinguish seemingly redundant concepts. In addition, the system automatically classifies terms when the definitions indicate that one term appears to have an 'is-a' relationship to another term. For example, since bacterial pneumonia is defined as occurring in the lung and caused by bacteria, it will automatically be subsumed by the concepts 'lung disease' and 'bacterial disease'.

Another research group, at the New Jersey Institute of Technology, has created an object-oriented schema for representing complex vocabularies. Using the MED, this research group was able to identify general 'area classes' of terms based on their definitional features [24]. The resulting simplified view allowed the MED content of 46 000 terms to be perceived as a much smaller set of 90 areas. As a result of this view, a small number of ambiguous terms were detected, based on the fact that their definitions contained the features of multiple, otherwise-mutually-exclusive areas.

2.3. Distributed vocabulary maintenance

The development and maintenance of a CMT typically occurs through one of two mechanisms: centralized (through one person) or distributed (through multiple members of an editorial committee). The centralized approach has the apparent advantage of consistency, but may suffer if the single person becomes a bottleneck in the process. The trade-off can be thought of 'too many cooks spoil the pot' versus 'many hands make light work'. In fact, neither approach is ideal. As a CMT grows large, it exceeds the mental capacity of a single person who will, however meticulous, eventually add new terms which are redundant with existing ones or add them in ways which are inconsistent with similar, previously added terms. On the other hand, members of a committee may act in an uncoordinated manner, even at cross purposes, and ultimately suffer from inefficiency or even become paralyzed.

Computer-based tools can help in both these approaches. To be 'smart', these tools must have knowledge about the definitions of the terms in the CMT. Consider, the following example, in which a person is attempting to add a new term to a CMT. In this case, the user has indicated that the new term is a disease, and the system knows that diseases often have features such as 'site' and 'etiology':

Computer:	Please enter the name of the new disease term.
Human:	Psittacosis.
Computer:	'Psittacosis' is a new disease name. Does Psittacosis have a site?
Human:	Yes, the lung.
Computer:	Does Psittacosis have an etiology?
Human:	Yes, *Chlamydia psittaci*.
Computer:	I already know about a disease which has the site 'lung' and etiology '*Chlamydia psittaci*'. It has the name 'Ornithosis'. Is 'Psittacosis' synonymous with 'Ornithosis'?
Human:	Yes.
Computer:	OK. I will add 'Psittacosis' as a synonym of the existing term 'Ornithosis'.

In this example, the computer-based tool has helped the human avoid the addition of a redundant term. Medical informatics researchers have theorized that such behavior is possible, if the CMT is modeled properly and contains definitional knowledge [7]. With tools such as this, consistency can be enforced by the system, rather than relying on one person to act in a consistent manner at all times or relying on the ability of a committee to be well-coordinated. With consistency enforced, the committee approach obtains a clear advantage over the single CMT author.

Researchers are beginning to explore the possibility of distributing the cognitive task of vocabulary development, using knowledge-based, artificially intelligent tools. The InterMed collaborative has explored the use of Ontolingua [23] for vocabulary modeling over the Internet [39]. The vocabulary server VOSER is being used at Intermountain Health Care to coordinate content development an institution-independent terminology across multiple hospitals [47].

Various application developers at the Presbyterian Hospital make use of a distributed vocabulary browser environment to explore the MED to determine what changes they need [1]. They then convert their set of changes into a standardized update format which is applied to the central vocabulary server. This process of browsing and updating makes little use of the knowledge in the MED to coordinate the cognitive work of the application developers. However, once the updates are applied to the MED, knowledge-based tools are used to audit the results [15]. Discrepancies are fed back to the authors of the offending updates, who are then responsible for making appropriate corrections. By this process, the MED editing tools attempt to enforce the agreed-upon cognitive model of the MED upon the various contributors.

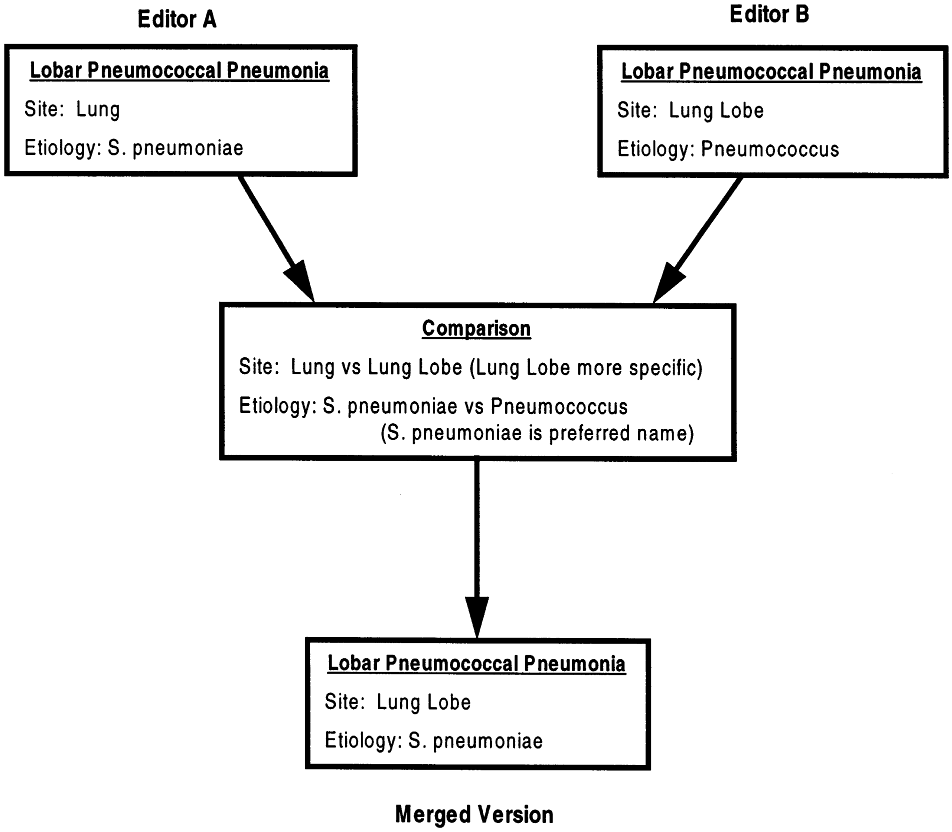

Fig. 2. This figure shows an example distributed vocabulary editing. In this example, two editors have created different definitions for the concept 'lobar pneumococcal pneumonia'. A formal comparison of these two description yields a set of differences which can be addressed through manual and/or automated means. For example, the resolution of two different values for the 'site' attribute might be accomplished with a rule that states "Always choose the more specific value", while the resolution of two different values for the 'etiology' attribute might be accomplished by mutual agreement of the editors. Although not shown in this figure, the resolution of the differences can be transmitted back to the editors individual version of the CMT to update each automatically such that it conforms with the merged version [8].

A large knowledge-supported effort for distributed terminology development is the Convergent Medical Terminology project, being conducted by Kaiser Permanente and the Mayo Clinic using SNOMED [8]. This work is providing valuable insights in how knowledge can be used to detect conflicts among terminology authors, which is a prerequisite to being able to understand why they arise and how they can be resolved (Fig. 2). Recently, for example, vocabulary editors from three different regions of the Kaiser Permanente organization used knowledge-based tools to create a set of 54 326 changes to their CMT. Many of the changes were in overlapping domains and the system was able to identify 1216 conflicts 2.2%. The system was further able to classify the conflicts into those which could be resolved automatically and those which required careful human review [9]. This experience not only demonstrates the value of the knowledge-based vocabulary for assisting in maintenance of the integrity of the CMT, but the low rate of conflicts indicates that the intelligent tools are valuable for coordinating disparate, remote editors such that errors are prevented (an inexpensive process) as well as corrected (a more expensive process).

3. Using knowledge-based CMTs to support distributed cognition in patient care

3.1. Defining distributed cognition in patient care

One of the primary purposes of CMTs is to support patient care applications. One of the primary reasons for having a standardized CMT is to enable data sharing and coordination of multiple applications. Applications can be typically classified as either primary clinical systems (those which record, store and present patient-specific information) and general medical knowledge sources (those which provide access to general information for use in solving patient-specific problems). Broadly defined as decision support systems (DSSs), examples of the latter include expert systems, rule-based alerting/reminder systems, and bibliographic retrieval systems.

DSSs are typically used by clinicians who identify a problem to be solved, select an appropriate application, interact with it directly, obtain the information they need and then act on it—a clear example of distributed cognition in a specific context, as defined at the beginning of this paper. A somewhat different approach is to integrate DSSs directly with clinical systems such that relevant patient information is transferred to the DSS directly. This is possible because the clinical system defines the context (e.g. a specific patient's laboratory results). If the information transfer is effective, then we can expect that this process will be a relatively more efficient way to distribute cognition, since the user is spared the task of transferring the information and perhaps even the need to interact directly with the DSS.

The integration of DSSs with clinical systems can take many forms, ranging from those in which the human user is responsible for data transfer, to those in which the transfer, and even the application selection, is performed automatically [17]. In these latter systems, a standardized CMT is essential for assuring that information collected about the patient for record-keeping purposes is properly represented in the DSS.

To date, the major successful integration efforts have involved the incorporation of rule-based reminder/alerting systems into hospital information systems [27,34,43]. In all these cases, the CMT was created for the record keeping purposes and the rules were written to use the same CMT. However, each relies on the availability of adequate resources for DSS development at the home institution and fails to take advantage of applications developed at other sites, using different CMTs.

3.2. Facilitating distributed cognition with CMTs

Some success has been achieved with DSS integration by translating patient data, recorded with one CMT, into DSS-specific terms [10,25,26,31,36,40,41]. Most of these systems utilize the UMLS to assist with the translation. Fig. 3 shows a typical scenario in which this kind of integration is achieved and the central role played by the CMT.

At Columbia University, a pilot system was developed for searching the bibliographic database Medline using generic questions that could be asked about patients' diagnoses and procedures [14]. For example, if a patient had previously been diagnosed as having had a myocardial infarction and had also undergone cardiopulmonary resuscitation, the clinical database would contain this information encoded with the ICD9-CM as:

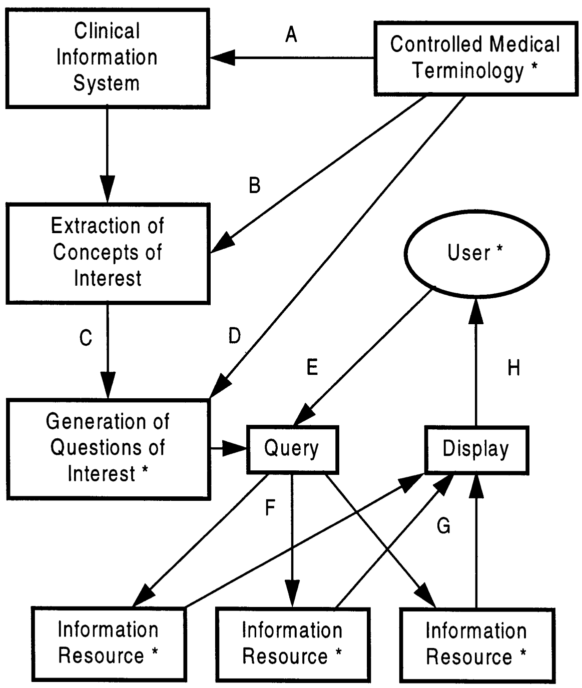

Fig. 3. Example of distributed medical decision support through the use of a controlled medical terminology (CMT). In this example, the CMT used to code clinical information (A) is also used to identify important medical concepts appearing in the electronic medical record (B) which, in turn, can be used for the generation of questions of potential interest to the user (C). The CMT plays a role in question generation by providing a means to translate clinical terms into those used by information resources (D). The user selects a question of interest (E) and an appropriate query is sent to the relevant resource (F). The result of the query is obtained (G) and displayed to the user (H). The items marked with an asterix indicate places where cognitive processes, whether human or artificial, are carried out.

410.81 Acute myocardial infarction of other specified sites, initial episode of care
99.60 Cardiopulmonary resuscitation, not otherwise specified

Using the UMLS, the system could translate these to MeSH keywords (used to index bibliographic citations in Medline) as:

Myocardial infarction
Cardiopulmonary resuscitation

The system would then generate questions concerning a disease and a procedure. In this case, the questions would be:

Is myocardial infarction treated by cardiopulmonary resuscitation?
Is myocardial infarction caused by cardiopulmonary resuscitation?
Is myocardial infarction diagnosed by cardiopulmonary resuscitation?
Is myocardial infarction prevented by cardiopulmonary resuscitation?

Each of these questions could be converted to a Medline query which could then be passed to Medline and the results of the query could be displayed to the user.

This 'Medline button' was technically feasible except that the translation from ICD9-CM to MeSH could only be accomplished for a minority of ICD9-CM terms. For those which could be translated, the translation was often found to be inappropriate for use in practical Medline searches. This, then, is an example of how distributed cognition did not achieve any efficiency over the traditional, manual method of literature retrieval.

3.3. Advantages of knowledge-based CMTs

In order to improve our ability to use outside information resources as DSSs, we have begun to explore ways in which we can make use of the knowledge in the MED. Because of the interrelationships among terms in the MED ('is-a' and nonhierarchical semantic links), terms found in a patient's medical record can be used to suggest other terms, more appropriate for use with a DSS.

For example, we wished to create a 'Medline button' which could be driven by laboratory test results. Although the tests are coded as controlled terms in the MED, translating them to MeSH would not generally be of help for Medline searches. If a patient has an elevated serum calcium level, for example, translating 'serum calcium measurement' to MeSH would allow us to identify journal articles about how calcium tests are performed, but not about how to treat an elevated calcium level. In order to obtain a more appropriate MeSH term, we query the MED to find out what substance serum calcium test measures and what specimen is used. The results, 'calcium' and 'serum', respectively, are recognized (using the UMLS) as valid MeSH terms, which can be used in the question 'How is elevated calcium in the serum treated?' Here, the knowledge in the MED is used (in a cognitive, albeit simple, task) to improve the quality of the information to be transferred and thus provide the potential for more efficient distribution of the original cognitive tasks-obtaining relevant medical literature citations.

In one implementation, we used knowledge about how tests (such as serum calcium test) relate to patient findings (such as hypercalcemia and hypocalcemia) to provide input to the expert system DXplain [2], available over the Internet. The clinical information system contains information about the tests patients have had and the numeric results of those tests, but does not necessarily contain the conceptual findings. DXplain takes these findings as input but can not interpret a numeric result of a test. Using the MED, however, we can easily convert an elevated calcium result to the DXplain term 'hypercalcemia'. Using this method, we are able to convert entire sets of test results (such as all the results of a 20-component chemistry panel) into DXplain findings which can then be passed to DXplain to query for a differential diagnosis [19]. This is a different example of distributed cognition: the MED serves as a resource for a system charged with the translation task, while DXplain carries out the differential diagnosis task.

Finally, we can use the knowledge of terms to help drive the selection of information resources. For example, when displaying laboratory results on in a world wide web-based application, we can detect whether any of the tests have, in the MED, a 'substance measured' relationship to 'cholesterol'. If true, the application which is responsible for creating the display can include a 'cholesterol guideline' button which the user may select. The button, in turn, triggers a program which converts the cholesterol tests result, along with other relevant patient data, into a form usable by a guideline processor [16]. The processor then uses the data to carry out the National Cholesterol Education Program's recommendation for the management of elevated cholesterol [37].

4. Discussion

In one sense, controlled medical terminologies have always supported distributed cognition for as long as they have been used for standardizing the exchange of information among care givers. Limitations on the quality of the terminologies has impaired the quality of information exchange and hence the effectiveness of distributed cognition. However, the evolution of terminologies into formal knowledge bases appears to support improved information quality. Terminology developers are in effect relegating some of the cognitive responsibilities to the CMTs themselves. As a result, application developers can rely on the CMT to support a certain amount of inferencing necessary for sophisticated activities, such as linking clinical systems to on-line information sources.

The most demanding responsibility will be the support of CMT development itself, since this will require the CMT to have internal consistency. Faulty inferencing applied to external tasks may produce bizarre results, but should at least be consistent, detectable, and (one hopes) correctable. Faulty inferencing applied to internal maintenance, on the other hand, may alter the CMT in nonmonotonic ways which may be hard to detect and could very possibly tend towards chaotic solutions. Fortunately, current experience with knowledge-based terminology maintenance is showing that restricting maintenance tasks to well-defined domains and monitoring the process closely leads to significant computer-assisted improvements in CMTs.

The successes with support of distributed cognition through knowledge-based CMTs is encouraging, but we are a long way off from having an intelligent CMT which can automatically exchange information among health care systems and, while doing this, maintain itself. A great deal more work is needed in the development of standards for terminological work-not just at the level of lists of terms, but in defining the representational structures needed for exchanging terms among systems. For example, every few years, someone demonstrates that frame-based representations of terms are useful for translation among different terminologies [11,32,46], but relatively little work has been done to develop consensus of what kinds of information these frames should represent (e.g. what slots they should contain) [20]. Only recently has a consensus emerged for some relatively small domains: diseases [9] and laboratory tests [21].

The real challenge for development of knowledge-based CMTs will be to provide explicit definitions for terms which are independent of external contexts of the terms. Explicit definitions will be needed if we are to develop computer-based tools which can help us maintain and use CMTs properly. While terms will almost never be used in a context-free way, they must have meanings which are stable across contexts if they are to be used by multiple systems to exchange patient information from one context to another.

As described in this paper, controlled medical terminologies can be both the target of, and resource for, distributed cognition. The intellectual effort invested in a terminology's knowledge base can empower intelligent tools that address the quality of the terminology. This improved quality, in turn facilitates the much larger task of coordinating health care practitioners and computer systems to provide coherent, relevant information about the care of specific patients.

References

[1] Allen, B.A., Barrows, R.C., Desai, N., AcessMED, a tool for browsing, searching and maintaining a large controlled medical vocabulary, in: Cimino, J.J. (Ed.), Proceedings of the 1996 AMIA Fall Symposium, Hanley and Belfus, Washington DC, 940 pp.

[2] G.O. Barnett, J.J. Cimino, J.A. Hupp, E.P. Hoffer, DXplain—An evolving diagnostic decision-support system, J. Am. Med. Assoc. 258 (1987) 67–74.

[3] C.E. Barr, H.J. Komorowski, E. Pattison-Gordon, R.A. Greenes, Conceptual modeling for the Unified Medical Language System, in: R.A. Greenes (Ed.), Proceedings of the Twelfth Annual Symposium on Computer Applications in Medical Care; Washington, Silver Spring (MD): IEEE Computer Society Press, 1988, pp. 148–51.

[4] Bernauer J., Conceptual graphs as a operational model for descriptive findings, in: P.D. Clayton (Ed.), Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington DC, November, 1991, pp. 214–218.

[5] P.J.B. Brown, M. O'Neil, C. Price, Semantic representation of disorders in version 3 of the Read Codes. Methods Inform. Med. (1988) in press.

[6] K.E. Campbell, M.A. Musen, Representation of clinical data using SNOMED III and conceptual graphs, in: Safran, C. (Ed.), Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care, Washington DC, McGraw-Hill, New York, 1993, pp. 354–358.

[7] K.E. Campbell, A.K. Das, M.A. Musen, A logical foundation for representation of clinical data, J. Am. Med. Inform. Assoc. 1 (1994) 218–232.
[8] K.E. Campbell, S.P. Cohn, C.G. Chute, G. Rennels, E.H. Shortliffe, Gálapagós: computer-based support for evolution of a convergent medical terminology, J. Am. Med. Inform. Assoc. (suppl) 3 (1996) 269–273.
[9] K.E. Campbell, Distrbuted development of a Logic-Based Controlled Medical Terminology. Ph.D. Dissertation, Stanford University, CA, June, 1997.
[10] C. Cimino, G.O. Barnett, L. Hassan, D.R. Blewett, J.L. Piggins, Interactive query workstation: standardizing access to computer-based medical resources, Comput. Methods Prog. Biomed. 35 (4) (1991) 293–299.
[11] Cimino, J.J., Barnett, G.O., Automated translation between medical terminologies using semantic definitions, In: L.W. Kingsland (Ed.), Proceedings of the American Association for Medical Systems and Informatics Congress, May 10, 1989, pp. 113–117. Reprinted in MD Comput. 72 (1990) 104–109.
[12] J.J. Cimino, G. Hripcsak, S.B. Johnson, P.D. Clayton, Designing an introspective, controlled medical vocabulary, in: L.W. Kingsland (Ed.), Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care, Washington DC, IEEE Computer Society Press, New York, 1989, pp. 513–518.
[13] J.J. Cimino, G. Hripcsak, S.B. Johnson, C. Friedman, D.J. Fink, P.D. Clayton, UMLS as knowledge base—a rule-based expert system approach to controlled medical vocabulary management, in: R.A. Miller (Ed.), Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care, 1990, Nov 4–7, Washington, Los Alamitos, CA, IEEE Computer Society Press, New York, 1990 pp. 175–179.
[14] J.J. Cimino, S.B. Johnson, A. Aguirre, N. Roderer, P.D. Clayton, The Medline Button, in: M.E. Frisse (Ed.), Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care, 1992, Baltimore, MD, McGraw-Hill, New York, 1992, pp. 81–85.
[15] J.J. Cimino, S.B. Johnson, G. Hripcsak, C.L. Hill, P.D., Clayton. Managing vocabulary for a centralized clinical system, in: S. Kaihara, R.A. Greenes (Eds.), Proceedings of the World Congress on Medical Informatics-Medinfo '95; Vancouver, Canada, Healthcare Computing and Communications Canada, Edmonton, Alberta, 1995, pp. 117–120.
[16] J.J. Cimino, S.A. Socratous, P.D. Clayton, Automated guidelines implemented via the World Wide Web (Poster), in: R.M. Gardner (Ed.), Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care; New Orleans, LA; October–November, Hanley and Belfus, Philadelphia, 1995, 941 pp.
[17] J.J. Cimino, Linking patient information systems to bibliographic resources, Methods Inform. Med. 35 (2) (1996) 122–126.
[18] R.A. Côté, D.J. Rothwell, J.L. Palotay, R.S. Beckett, L. Brochu (Eds), The Systematized Nomenclature of Medicine: SNOMED International, College of American Pathologists, Northfield, Illinois, 1993.
[19] G. Elhanan, S.A. Socratous, J.J. Cimino, Integrating DXplain into a clinical information system using the world wide web, in: J.J. Cimino (Ed.), Proceedings of the American Medical Informatics Association Annual Fall Symposium (formerly SCAMC); Washington, DC; October, Hanley and Belfus, Philadelphia, 1996, pp. 348–352.
[20] C. Friedman, S.M. Huff, W.R. Hersh, E. Pattison-Gordon, J.J. Cimino, The Canon effort: working toward a merged model. J. Am. Med. Inform. Assoc. (1995) 4–18.
[21] A.W. Forrey, C.J. McDonald, G. DeMoor, S.M. Huff, D. Leaville, D. Leland, T. Fiers, L. Charlse, B. Griffin, F. Stalling, A. Tullis, K. Hutchins, J. Baenziger, Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results, Clin. Chem. 42 (1996) 81–90.
[22] E.R. Gabrieli, A new electronic medical nomenclature, J. Med. Sys. 13 (6) (1989) 355–373.
[23] T. Gruber, Ontolingua: A mechanism to support portable ontologies, Technical Report No. 91–66, Knowledge Systems Laboratory, Stanford University, CA, 1992.
[24] H. Gu, M. Halper, J. Geller, J.J. Cimino, Y. Perl, Utilizing OODB schema modeling for vocabulary management, J. Am. Med. Inform. Assoc. (suppl) 3 (1996) 274–278.

[25] J.W. Hales, R.C. Low, K.T. Fitzpatrick, Using the Internet Gopher protocol to link a computerized patient record and distributed electronic resources, in: C. Safran (Ed.), Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care, Washington, DC, November, McGraw-Hill, New York, 1993, pp. 621–625.
[26] J.E. Hammond, W.E. Hammond, W.W. Stead, Information management through integration of distributed resources: the TMR-NLM connection and prototype, in: R.A. Miller (Ed.), Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC, November, 1990, pp. 719–23.
[27] G. Hripcsak, P.D. Clayton, R.A. Jenders, J.J. Cimino, S.B. Johnson, Design of a clinical event monitor, Comput. Biomed. Res. 29 (3) (1996) 194–221.
[28] S. Johnson, C. Friedman, J.J. Cimino, Clark T.G. Hripcsak, P.D. Clayton, Conceptual data model for a central patient database, in: P.D. Clayton (Ed.), Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC, November, 1991, pp. 381–385.
[29] Lambert, H., Wood, M. (Eds.), International Classification of Primary Care, Oxford University Press, Oxford, 1987.
[30] D.A.B. Lindberg, B.L. Humphreys, A.T. McCray, The Unified Medical Language System, Methods Inform. Med. 32 (1993) 281–291.
[31] J.W. Loonsk, R. Lively, E. TinHan, H. Litt, Implementing the Medical Desktop: tools for the integration of independent information resources, in: P.D. Clayton (Ed.), Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington DC, November, 1991, pp. 574–7.
[32] F.E. Masarie Jr., R.A. Miller, O. Bouhaddou, N.B. Giuse, H.R. Warner, An interlingua for electronic interchange of medical information: using frames to map between clinical vocabularies, Comput. Biomed. Res. 24 (4) (1991) 379–400.
[33] E.K. Mays, R.A. Weida, R.A. Dionne, M. Laker, B.F. White, C. Liang, F.J. Oles, Scalable and expressive medical terminologies, J. Am. Med. Inform. Assoc. (suppl) 3 (1996) 259–263.
[34] C.J. McDonald, W.M. Tierney, J.M. Overhage, D.K. Martin, G.A. Wilson, The Regestrief medical record system: 20 years of experience in hospitals, clinics, and neighborhood health centers, MD Comput. 9 (24) (1992) 206–217.
[35] R.A. Miller, M.A. McNeil, S.M. Challinor, F.E. Masarie, J.D. Myers, The Quick Medical Reference project-status report, West. J. Med. 145 (1986) 816–822.
[36] R.A. Miller, F.M. Gieszczykiewicz, J.K. Vries, G.F. Cooper, CHARTLINE: providing bibliographic references relevant to patient charts using the UMLS metathesaurus knowledge sources, in: M.E. Frisse (Ed.), Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care, Baltimore, MD, November, McGraw-Hill, New York, 1992, pp. 86–90.
[37] National Cholesterol Education Program (NCEP), Summary of second report by the expert panel on detection, evaluation and treatment of high blood pressure in adults, Adult Treatment Panel II, J. Am. Med. Assoc. 269 (23) (1993) 3015–3023.
[38] National Library of Medicine, Medical Subject Headings, NTIS NLM-MED-92–01, The Library, Bethesda, MD, (annual).
[39] D.E. Oliver, M.R. Barnes, G.O. Barnett, H.C. Chueh, J.J. Cimino, P.D. Clayton, W.M. Detmer, J.H. Gennari, R.A. Greenes, S.M. Huff, M.A. Musen, E. Pattison-Gordon, E.H. Shortliffe, S.A., Socratous, S.W. Tu, InterMed: an Internet-based medical collaboratory, in: R.M. Gardner (Ed.), Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care, New Orleans, LA, Hanley and Belfus, Philadelphia, 1995, 1023 pp.
[40] S.M. Powsner, C.A. Riely, K.M. Barwick, J.S. Morrow, P.L. Miller, Automated bibliographic retrieval based on current topics in hepatology: Hepatopix, Comput. Biomed. Res. 22 (1989) 552–564.
[41] S.M. Powsner, P.L. Miller, From patient records to bibliographic retrieval: a Meta-1 front-end, in: P.D. Clayton (Ed.), Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington, DC, 1991, pp. 526–530.
[42] C. Price, M. O'Neil, T.E. Bentley, P.J.B. Brown, Exploring the ontology of surgical procedures in the Read Thesaurus, Methods Inform. Med. (1998) in press.

[43] T.A. Pryor, R.M. Gardner, P.D. Clayton, H.R. Warner, The HELP system., J. Med. Syst. 7 (2) (1983) 87–102.
[44] A.L. Rector, W.A. Nowlan, The GALEN project, Comput. Methods Prog. Biomed. 45 (1-2) (1994) 75–78.
[45] A.L. Rector, Thesauri and formal classifications: terminologies for people and machines, Methods Inform. Med. (1988) in press.
[46] R.A. Rocha, B.H. Rocha, S.M. Huff, Automated translation between medical vocabularies using a frame-based interlingua, in: P.D. Clayton (Ed.), Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, Washington DC, November, 1991, pp. 690–694.
[47] R.A. Rocha, S.M. Huff, P.J. Haug, H.R. Warner, Designing a controlled medical vocabulary server: the VOSER project, Comput. Biomed. Res. 27 (1994) 472–507.
[48] D.J. Rothwell, R.A. Côté, Managing information with SNOMED: understanding the model, J. Am. Med. Inform. Assoc. (suppl) 3 (1996) 80–83.
[49] United States National Center for Health Statistics, International Classification of Diseases, 9th ed., with Clinical Modifications, The Center, Washington, DC, 1980.
[50] J. van der Lei, J.S. Duisterhout, H.P. Westerhof, E. van der Does, P.V.M. Cromme, W.M. Boon, J.H. van Bemmel, The introduction of computer-based patient records in The Netherlands, Ann. Intern. Med. 119 (10) (1993) 1036–1041.
[51] E.T. Wong, T.A. Pryor, S.M. Huff, P.J. Haug, H.R. Warner, Interfacing a stand-alone diagnostic expert system with a hospital information system, Comput. Biomed. Res. 27 (1994) 116–129.