Automatic Structuring of Radiology Free-Text Reports[1]

Ricky K. Taira, PhD • Stephen G. Soderland, PhD • Rex M. Jakobovits, PhD

A natural language processor was developed that automatically structures the important medical information (eg, the existence, properties, location, and diagnostic interpretation of findings) contained in a radiology free-text document as a formal information model that can be interpreted by a computer program. The input to the system is a free-text report from a radiologic study. The system requires no reporting style changes on the part of the radiologist. Statistical and machine learning methods are used extensively throughout the system. A graphical user interface has been developed that allows the creation of hand-tagged training examples. Various aspects of the difficult problem of implementing an automated structured reporting system have been addressed, and the relevant technology is progressing well. Extensible Markup Language is emerging as the preferred syntactic standard for representing and distributing these structured reports within a clinical environment. Early successes hold out hope that similar statistically based models of language will allow deep understanding of textual reports. The success of these statistical methods will depend on the availability of large numbers of high-quality training examples for each radiologic subdomain. The acceptability of automated structured reporting systems will ultimately depend on the results of comprehensive evaluations.

Abbreviations: ASCII = American Standard Code for Information Interchange, UMLS = Unified Medical Language System, XML = Extensible Markup Language

Index terms: Computers • Computers, diagnostic aid • Radiology and radiologists, design of radiological facilities • Radiology reporting systems

Radiology 2001; 21:237–245

[1]From the Department of Radiology, Children's Hospital and Regional Medical Center, 4800 Sandpoint Way NE, Mailstop CH-69, Seattle, WA 98105 (R.K.T.); and the Department of Radiological Sciences, University of California, Los Angeles (S.G.S., R.M.J.). Presented as an *info*RAD exhibit at the 1999 RSNA scientific assembly. Received March 24, 2000; revision requested May 1 and received June 1; accepted June 13. Supported in part by National Institutes of Health Grant No. P01-CA51198. Address correspondence to R.K.T. (e-mail: *rtaira@u.washington.edu*).

```
Topic:  'Mass'      Object Instance ID:  xxxxyyy

• has existence      Value              = TRUE
                     Certainty          = DEFINITE
                     How Determined     = OBSERVATION
                     When               = CURRENT_EXAM

• has location:      Value              = 'right lower lobe'
                     Relation           = 'in'
                     When               = CURRENT_EXAM

• has size:          Value              = '5cm'
                     Precision          = approximately
                     Relation           = EQUALS
                     Dimension          = 'maximum diameter'
                     When               = CURRENT_EXAM

• has size trend:    Value              = 'unchanged'
                     Relation           = EQUALS
                     Reference Event    = 'previous examination'
```

Figure 1. Output knowledge frame for the sample sentence, "A mass is seen in the right lower lobe that measures 5 cm in maximum diameter and is unchanged from the previous examination." Values from a controlled vocabulary are shown in capital letters, and values as expressed in the input text are enclosed in single quotation marks.

Introduction

Radiology reports contain a great deal of information that characterizes a patient's medical condition. However, a large percentage of this information is unstructured, taking the form of free text, and is therefore difficult to search, sort, analyze, summarize, and present.

Previous studies have demonstrated the potential benefits of structured medical data for medical practice, research, and teaching. In the clinical setting, structured reports can be used to help organize and improve the presentation of the medical record (1,2). For example, if a radiologist is interested only in a given clinical episode, he or she may elect to retrieve only those reports in which the relevant anatomy or findings are described. Accurate extraction of lesion size information from radiology reports can allow a system to automatically construct a growth timeline for an indicator lesion. Accurate extraction of imaging section information (eg, "mass seen on cut #12") can allow default presentation of key images. A clinical expert system can use structured report information for decision support (3,4) and to automatically flag alarm conditions such as pneumothorax (5,6). For research and teaching, structured reports can greatly improve the recall and precision of information retrieval tasks. Only structured data are amenable to advanced causal, spatial, temporal, and evolutionary database modeling techniques that are now being developed in the fields of medical informatics and computer science (7–9). Complex queries such as "Find 50 cases of spiculated masses greater than 5 cm in diameter in the left upper lobe" can be expected to yield the desired results. The implications for teaching files and data collection for retrospective research studies are obvious. Finally, the most practical means of indexing radiologic image data is by extracting the important findings from the associated dictated report. Thus, structured reports serve as a key for creating medical multimedia digital libraries.

However, automatic structuring of radiology reports is a difficult task for the following reasons:

1. Automatic structuring requires deep understanding because it is desirable to translate all relevant information into structured form.

2. Automatic structuring must deal with ungrammatical writing styles. Shorthand and telegraphic writing styles are comemon in radiology reports. A natural language understanding system that works for the *New York Times* will not work in the medical domain. Each subspecialty of radiology may have different language models (in the statistical sense). These differences are especially noticeable when comparing interventional and purely diagnostic reports. In addition, there are many stylistic variations between radiologists, and long, run-on sentences are not uncommon. Negations and conjunctive lists (eg, *A*, *B*, *C*...) are also common in radiology reports.

3. The vocabulary is large. Large numbers of complex medical terms, proper names, product names, abbreviations, and staging codes are used in radiology reports. Hundreds of descriptive adjectives are used that are not found in any common electronic medical glossaries.

4. There is an assumed knowledge between the writer and reader. The radiologist understands the

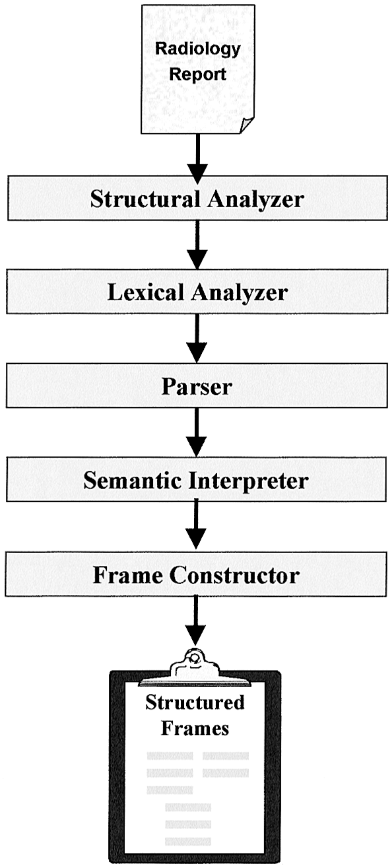

Figure 2. Diagram illustrates the processing architecture of our natural language processing system.

world of the referring physician and vice versa. Consequently, details are often left out because they are assumed to be common knowledge.

The task of automatically structuring radiology reports can be divided into three subtasks. The first is creating expert-defined data models of the targeted information contained within the report (10–15). The second subtask is creating a medical natural language processor that can map the free-text report dictated by the radiologist into the knowledge structures defined in the first subtask. Several medical natural language processing systems are currently being researched and are in various stages of clinical testing (3,6,16–19). Spyns (20) has provided an excellent review article on medical natural language processing systems. The third subtask is mapping the slot values within the output knowledge frames into an understandable controlled vocabulary (eg, Systematized Nomenclature of Medicine [SNOMED], Unified Medical Language System [UMLS]) (16,17,21,22).

We have developed a system that automatically structures the important medical information contained in a radiology free-text document as a formal information model that can be interpreted by a computer program. The input to the system is an American Standard Code for Information Interchange (ASCII) free-text report from a radiologic study. The system imposes no constraints on the dictation style of the radiologist. The system outputs a formalized representation of the important information contained within the free-text document in the form of a structured report (23,24). This includes information on the existence, properties, location, and diagnostic interpretation of findings. For example, consider the following sentence: "A mass is seen in the right lower lobe measuring 5 cm in maximum diameter and is unchanged from the previous examination." The desired output frame for this sample sentence is shown in Figure 1. Note that the frame consists of a head (or topic) and a set of property subframes. Each property subframe contains a value and any expert-defined context modifiers for that property value (eg, time of measurement, precision, dimension). Frame slots can contain either values from a controlled vocabulary or values as expressed in the input text.

In this article, we focus mainly on the second subtask involved in automatically structuring radiology reports (ie, natural language processing). We discuss our system in terms of overall architecture (structural analyzer, lexical analyzer, parser, semantic interpreter, frame constructor), the current status of our project, and possible future trends.

System Architecture

Figure 2 shows the overall architecture of our natural language processing system. Our system is functionally similar to other medical natural language processing systems and consists of the following collaborative modules:

1. The *structural analyzer* isolates the sections of radiology reports (eg, "Procedure Description," "History," "Findings," "Impressions") and the individual sentences within these sections.

2. The *lexical analyzer* looks up semantic and syntactic features of words with use of a medical lexicon.

3. The *parser* determines the modifier-head relations between words in a sentence.

4. The *semantic interpreter* interprets the links of the parser-generated dependency diagram and outputs a set of logical relations.

5. The *frame constructor* bundles the individual logical relations into structured frames.

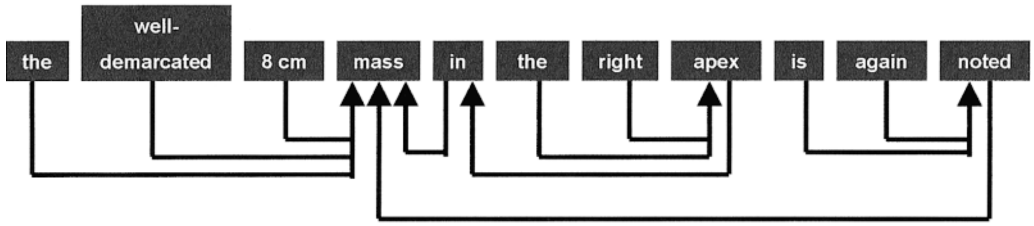

Figure 3. Dependency diagram generated by the parser shows modifier-head relations between words. The tail of an arc represents the modifier, whereas the tip represents the head.

Structural Analyzer

The input to the structural analysis module is a free-text report received from a radiology information system. The structural analyzer performs a preprocessing step that standardizes the character representation of the text and removes custom text formatting. We convert all text to International Organization for Standardization (ISO) ASCII standard representation. Reports at our institution, for example, are stored using a DOS-based character set. The effect is seen mainly in the representation of control characters such as line feeds. In older radiology information systems, it is common for the text to be custom formatted for VT-100–type terminals and for the text to contain long sequences of blanks between words. We use simple heuristic rules to remove these unnecessary white spaces.

After this preprocessing step, the structural analyzer identifies the various sections of the report. A section is informally defined as follows: If the report is viewed as a document form, the sections correspond to the fields of the form. Often, there are certain major sections that are consistently seen (eg, "Study Description," "Clinical History," "Findings," "Conclusions") as well as context-dependent or minor sections (eg, "Age," "Sedation," "Social History," "Cardiac Status"). Major sections usually begin on a new line. Minor sections may or may not do so, depending on the length of the section body. The structural analyzer uses a simple construction model of a radiology report and a rule-based classifier to detect section boundaries. The model includes a list of common section labels (eg, "Clinical History," "Findings," "Impressions") and common formatting layouts (eg, section order, section length, use of colons).

Once the text within each individual section has been determined, the next step is to identify the individual sentences within each section. The algorithm for determining sentence boundaries uses a maximum entropy classifier (25). The classifier in turn uses some 40 overlapping features to determine whether a period corresponds to an end-of-sentence marker. The system was trained on 6,500 sentences from a test corpus of radiology reports.

Lexical Analyzer

The input to the lexical analyzer is a single sentence, and the lexical analyzer outputs the semantic and syntactic features of each word or phrase in the sentence. In this labeling process, punctuation is identified; dates, numeric measurements, special symbols (eg, TNM cancer staging codes), and proper nouns are recognized with use of special-purpose, finite-state machines; words and phrases are looked up in medical lexicons; and prefixes and suffixes are analyzed. Any words that remain unknown after this process are inserted into a database of unrecognized words together with their frequency of occurrence. A medical language expert is responsible for later assigning semantic and syntactic features to these unknown words.

Our lexicons were developed manually. We maintain both single words and multiple-word phrases within our lexicon. The latter are useful when they represent a concept that cannot be precisely described using only the component words. We do not use existing lexical sources such as the UMLS (24,26) because we believe that they do not currently support the level of knowledge or detail required for a natural language understanding system. (For example, they do not contain a sufficient number of semantic classes to support our statistical parsing and semantic interpretation algorithms. In addition, the coverage of important descriptive adjectives common in radiology is as yet insufficient.) Words and word phrases to be included in the lexicon were gathered from two distinct types of sources: published sources and actual radiology reports.

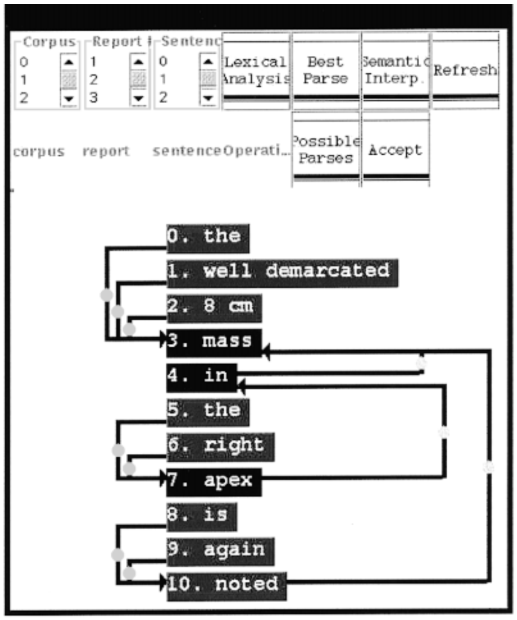

Figure 4. The semantic interpreter derives the logical relations based on the parser-generated dependency diagram (cf Fig 3). The links between concepts 3, 4, and 7 indicate the logical relation "has-location." The arguments (ie, head, relationship between head and value, and modifying value) are "mass," "in," and "apex," respectively.

The published sources included indexes from thoracic radiology textbooks (27), radiology review manuals (28), radiology word compilations (29), and published thoracic radiology glossaries (30,31). The terms from these published sources ensure the generality of the concepts covered by the lexicon but do not include all the string representations for these concepts. The collection of words and phrases from actual radiology reports ensures that the system works at a practical level and that most of the string representations for the basic concepts related to the target findings are included. Our current radiology lexicon has been developed for the fields of thoracic radiology and neuroradiology and includes over 5,000 entries with over 200 semantic classes. Each entry in the lexicon is characterized by a concept identifier and a set of features organized into multiple channels. Within each channel, features are organized hierarchically, from general to specific.

Parser

The input to the parser is the semantic and syntactic features of each word or phrase in the sentence, and the parser outputs a dependency diagram showing relations between the words of the sentence (Fig 3). An arc from word A to word B indicates that word A modifies word B. The arcs are left unlabeled, with the modifier semantics left for the following step (semantic interpretation). The parser algorithm makes use of statistical methods and relies on two sets of probabilities: (*a*) the "affinity" between words A and B, and (*b*) the "valence" preferences of each word in the sentence. The dependency diagram that globally maximizes these probabilities is the one that is selected.

Word affinities describe the probability that word A modifies word B as seen in a set of training documents. Word affinity probabilities were estimated as follows: (*a*) A large sample of documents were collected from the domain of interest (thoracic radiology). (*b*) Training data were created by manually indicating the dependency diagram for each sentence to reflect the output from an ideal parser. Figure 4 shows the graphical interface used to create training examples. (*c*) For every pair of words (A, B) in each training sentence, statistics were collected on how often word A modified word B. In addition, their relative sentence order and the "distance" between them were recorded. The result of the training set was a table containing statistics for each pair of words (A, B) in all the contexts that were encountered.

Estimating word affinity statistics from a finite set of training data leads to the problem of sparse data. The number of combinations of all possible words A and B that could occur in the same sentence is extremely large. A new sentence is likely to have pairs of words that have never been encountered in a training set. Two methods of "smoothing" the statistics are used: (*a*) Estimating statistics from the semantic and syntactic features of the words rather than the words themselves. The number of unique semantic classes is on the order of a few hundred, whereas the number of syntactic classes is 15. (*b*) Hierarchically organizing the features themselves and generalizing them by dropping the most specific features. For example, the word "mass" has the semantic features *abnorm.physobj.finding.lesion*. The more general semantic class *abnorm.physobj.finding* covers more words, and *abnorm.physobj* covers an even larger set of words. This increased number of words can be used to improve the probability estimate of word-word affinities.

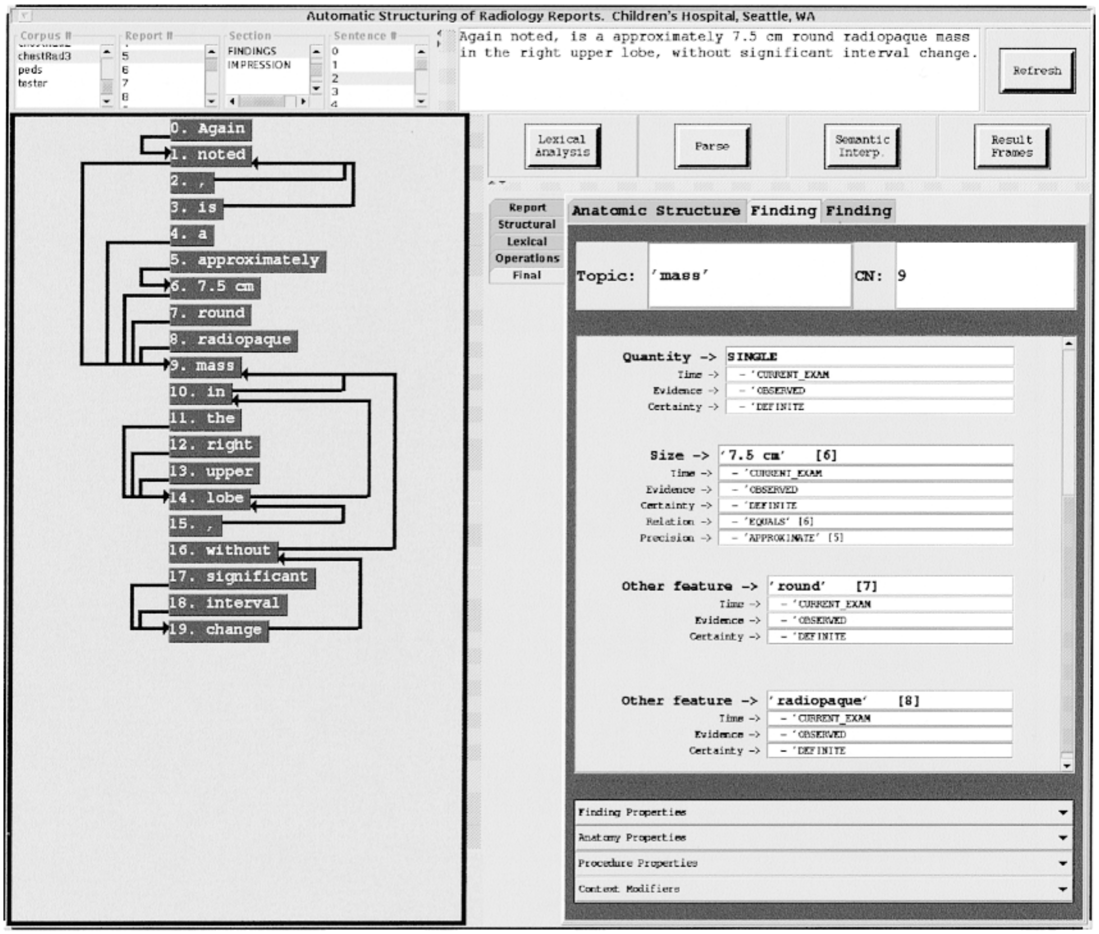

Figure 5. Screen shows a sample frame output (right) with intermediate parser results (left).

In addition to the affinity between pairs of words, our linguistic model considers the valence of individual words. This refers to the preference of words for certain types of complements and provides additional contexts to help determine whether a word A modifiers another word B. For example, a verb "prefers" to have one direct object rather than two. A noun that is the object of a preposition "prefers" not to have a verbal complement. We determine the valence probability of a word in a given parse structure from the number and type of arcs going to and from the word. Valence statistics are tabulated from the same set of annotated training sentences that are used for affinity statistics. In an earlier study, we provided additional details of the parser algorithm (32).

Semantic Interpreter

The input of the semantic interpreter is the dependency diagram of the parser. The semantic interpreter verifies that the arcs of the dependency diagram make medical sense and creates logical relationships that indicate the type of conceptual dependency between pairs of words. Each logical relation has a name (eg, "has-size") and three arguments. The first argument is the head, which is the word being modified (eg, "mass"). The second argument indicates the relationship between the head and the modifying value (eg, "in"), and the third argument is the modifying value itself (eg, "apex"). The logical relations serve to normalize alternate ways of expressing the same concept in free text (eg, "a right apical mass" and "a mass in the right apex" have identical logical relationships). Consider the dependency diagram for the sentence, "The well-demarcated 8-cm mass in the right apex is again noted" (Fig 4). The Table shows the logical relationships that can be inferred from this sentence and that together form a semantic network representing the sentence conceptually. In some cases, an argument value is selected from a predefined set of labels. The second

Logical Relations Inferred from a Sample Sentence*

Name of Logical Relation	Head	Argument Relation	Value
"has–external architecture"	'mass'†	EQUALS‡	'well-demarcated'
"has-size"	'mass'	EQUALS	'8-cm'
"has-location"	'mass'	'in'	'apex'
"has-direction"	'apex'	EQUALS	'right'
"has-evidence"	'mass'	EQUALS	OBSERVED
"has-newness"	'mass'	EQUALS	OLD

*"The well-demarcated 8-cm mass in the right apex is again noted."
†Values as expressed in the input text are enclosed in single quotation marks.
‡Values from a controlled vocabulary are shown in capital letters.

argument of the "Evidence" predicate is chosen from the set ("Observed," "Inferred") and is inferred from the link between "noted" and "mass." The predicate "has-newness" was created with the arc from "again" to "noted." Currently, there are 51 logical relations defined for the Findings section of radiology reports.

The semantic interpreter algorithm uses two techniques that can be used separately or in combination. The first technique makes use of symbolic rules that are learned automatically from a set of hand-tagged training examples. The system developer uses a graphical interface to indicate the logical relations associated with a training sentence. The system builds rules by comparing the parse diagram for a training sentence with a target logical relation to find a proposed rule that creates that logical relation. Some semantic features of the training sentence may not be essential to the rule. For example, if the training sentence has the word "mass," the rule does not necessarily require the semantic feature "lesion" to be present but might generalize to cover all words with the feature "finding." Generalizations of a proposed rule are tested on all other training sentences, and the generalization is selected that covers as many training sentences as possible without making errors. Before a rule is finally accepted, the human system developer is given a chance to edit it. The rule-learning algorithm is similar to the one used in File Oriented Interpretive Language (FOIL) (33) and Concurrent Representation of Your Space-Time Algorithms (CRYSTAL) (34). Currently, it includes over 450 machine-generated rules. In the second technique used by the semantic interpreter, a maximum entropy classifier is constructed for each logical relation (25,35).

Frame Constructor

The final processing step is to bundle all the logical relations that were found into appropriate frames (23). Each frame represents knowledge about a specific topic (eg, "mass"), together with descriptions of select properties. There are three classes of topics: (*a*) abnormal findings, (*b*) anatomy, and (*c*) medical procedures. For abnormal findings, there exists 11 types of properties: existence, location, quantity, size, severity, trend, normalcy, external architecture, interpretation, association, and "other." Each property subframe has relevant context modifiers such as time, evidence, certainty, degree, and dimension. The anatomy topic class includes subframes for normalcy, subparts, direction, and distribution modifiers. The medical procedure topic class includes subframes for reason for procedure, technical description, and anatomic site. Figure 5 shows a sample frame output along with intermediate parser results.

Current Status

The current project has been under development since 1995. We have concentrated mainly on thoracic radiology reports in lung cancer patients. Each of the five modules shown in Figure 2 is being developed as a separate server application that outputs results in Extensible Markup language (XML) format. Each module can run on separate computer processors. All code has been written in the JAVA programming language. A large proportion of our software consists of development tools for searching out and tagging training examples, applying classification algorithms to training data, and performing focused technical evaluations on isolated problem definitions (eg, lexical analysis, syntactic parsing).

We reported on a pilot technical evaluation at the 85th Scientific Assembly and Annual Meeting of the Radiological Society of North America in 1999. We conducted a tenfold cross-validation study and used 470 randomly selected sentences from the domain of thoracic radiology (from computed tomographic [CT] and projectional x-ray reports only). Our evaluation measures included recall (ie, the number of correct answers divided by the number of possible answers) and precision (ie, the number of correct answers divided by the number of reported answers). The results of the evaluation were as follows: For the parser, recall was 87% and precision was 88%. This evaluation included 4,314 possible parser links. For the semantic interpreter, recall was 79% and precision was 87%. The semantic interpreter evaluation included 4,300 possible semantic relations.

Our current clinical area of focus is serial CT reports obtained in pediatric patients with metastatic lung disease. The technical difficulty in this setting is tracking objects across sentence and report boundaries (ie, the linguistics-related problem of coreference resolution).

Discussion

What are the ultimate uses for a radiology natural language processor? For an ideal system, the uses presented in this article could be applicable. For any imperfect system, however, we must be cautious and aware of the types of errors or performance that can be expected. For example, a system that identifies true-positive results well but also reports some false-positive results might be used only as a front-end step for case searching followed by a manual review. A system that has low recall but high precision could be used for searching for teaching files. A system with both high recall and high precision could be used directly for population studies.

The accuracy and hence the utility of a medical natural language processor relies heavily on the number and diversity of training examples, and unfortunately, no large-scale training corpus for radiology reports is currently available. Furthermore, the accuracy of a language system is not uniform in terms of the specific information it extracts. For example, a system trained heavily on reports that contain many measurements may extract size information well but perform poorly in terms of location information. Fortunately, in radiology, we may be able to decide a priori what are the important types of information for a given type of study. Thus, information modeling of radiology reports is an important codevelopment. System development could then proceed by emphasizing training on high-priority information items (eg, size, location, trend).

Conclusions

The technology for implementing an automated structured reporting system is progressing well. Numerous investigators have addressed important aspects of this difficult problem, including researchers in medical knowledge representation, natural language processing, and systems engineering. The initial success of this system holds out hope that similar statistically based models of language will be accurate enough for deep understanding of textual reports. The success of these statistical methods will rely on the availability of a large number of high-quality training examples for each subdomain of radiology. The acceptability of automated structured reporting systems will ultimately depend on the results of comprehensive evaluations (36). Finally, XML is emerging as the preferred syntactic standard for representing and distributing these structured reports within a clinical environment (37–39).

References

1. Shortliffe EH, Hubbard SM. Information systems in oncology. In: De Vita VT, Hellman S, Rosenberg S, eds. Cancer: principles and practice of oncology. Philadelphia, Pa: Lippincott, 1989; 2403–2412.
2. Aberle DR, Dionisio JDN, Cardenas AF, et al. Integrated multimedia timeline of medical images and data for thoracic oncology patients. RadioGraphics 1996; 16:669–681.
3. Friedman C, Alderson P, Austin J, Cimino J, Johnson S. A general natural-language text processor for clinical radiology. J Am Med Inform Assoc 1994; 1:161–174.
4. Heckerman DE, Nathwani BN. Toward normative expert systems. II. Probability-based representations for efficient knowledge acquisition and inference. Methods Inf Med 1992; 31:106–116.
5. Lyman M, Sager N, Tick L, Nhan N, Borst F, Scherrer JR. The application of natural-language processing to healthcare quality assessment. Med Decis Making 1991; 11(suppl):65–68.

6. Sager N, Lyman M, Nhan NT, Tick L. Medical language processing: applications to patient data representation and automatic encoding. Methods Inf Med 1995; 34:140–146.
7. Muller R, Thews O, Rohrbach C, Sergl M, Pommerening K. A graph-grammar approach to represent causal, temporal and other contexts in an oncological patient record. Methods Inf Med 1996; 35:127–141.
8. Dionisio JDN, Cardenas AF. MQuery: a visual query language for multimedia, timeline, and simulation data. J Vis Lang Comput 1996; 7:377–401.
9. Chu WW, Hsu CC, Cardenas AF, Taira RK. Knowledge-based image retrieval with spatial and temporal constructs. IEEE Trans Knowledge Data Eng 1998; 10:872–888.
10. Bell DS, Pattison-Gordon E, Greenes RA. Experiments in concept modeling for radiographic image reports. J Am Med Inform Assoc 1994; 1:249–262.
11. Friedman C, Cimino JJ, Johnson SB. A schema for representing medical language applied to clinical radiology. J Am Med Inform Assoc 1994; 1:233–248.
12. Friedman C, Huff SM, Hersh WR, Pattison-Gordon E, Cimino JJ. The Canon Group's effort: working toward a merged model. J Am Med Inform Assoc 1995; 2:4–18.
13. Racha RA, Huff SM, Haug PJ, Evans DA, Bray BE. Evaluation of a semantic data model for chest radiology: application of a new methodology. Methods Inf Med 1998; 37:477–490.
14. Bidgood WD Jr. The SNOMED DICOM microglossary: controlled terminology resource for data interchange in biomedical imaging. Methods Inf Med 1998; 37:404–414.
15. Bidgood WD Jr, Bray B, Brown N, et al. Image acquisition context: procedure description attributes for clinically relevant indexing and selective retrieval of biomedical images. J Am Med Inform Assoc 1999; 6:61–75.
16. Haug PJ, Ranum DL, Frederick PR. Computerized extraction of coded findings from free-text radiologic reports. Radiology 1990; 174:543–548.
17. do Amaral MB, Satomura Y. Structuring medical information into a language-independent database. Med Inform 1994; 19:269–282.
18. Rassinoux AM, Baud R, Scherrer JR. Proximity processing of medical text. In: O'Moore R, Bengtsson S, Bryant SJ, Bryden J, eds. Medical Informatics Europe 90. Berlin, Germany: Springer-Verlag, 1990; 625–630.
19. Hripcsak G, Friedman C, Alderson PO, DuMouchel W, Johnson SB, Clayton PD. Unlocking clinical data from narrative reports: a study of natural language processing. Ann Intern Med 1995; 122:681–688.
20. Spyns P. Natural language processing in medicine: an overview. Methods Inf Med 1996; 35:285–301.
21. De Bruijn LM, Hasman A, Arends JW. Automatic SNOMED classification: a corpus-based method. Comput Methods Programs Biomed 1997; 54:115–122.
22. Friedman C. The UMLS coverage of clinical radiology. Proceedings of the Annual Symposium on Computer Applications in Medical Care. New York, NY: McGraw-Hill, 1992; 309–313.
23. Winston PH. Artificial intelligence. 2nd ed. Reading, Mass: Addison-Wesley, 1994; 251–290.
24. Bean CA. Formative evaluation of a frame-based model of locative relationships in human anatomy. Proc AMIA 1997: 625–629.
25. Ratnaparkhi A. Maximum entropy models for natural language ambiguity resolution. Dissertation. University of Pennsylvania, Philadelphia, 1998.
26. Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med 1993; 32:281–291.
27. Fraser RG, Pare JAP, Pare PD, Fraser RS, Genereux GP. Diagnosis of diseases of the chest. Vols I–II. Philadelphia, Pa: Sanders, 1991.
28. Dahnert W. The radiology review manual. 2nd ed. Baltimore, Md: Williams & Wilkins, 1993; 257–334.
29. Stedman's radiology words. Baltimore, Md: Williams & Wilkins, 1993.
30. Tuddenham WJ. Glossary of terms for thoracic radiology: recommendations of the Nomenclature Committee of the Fleischner Society. AJR Am J Roentgenol 1984; 143:509–517.
31. Webb WR, Muller NL, Naidich DP. High-resolution CT of the lung. New York, NY: Raven, 1992; 157–159.
32. Taira RK, Soderland SG. A statistical natural language processor for medical reports. In: Proceedings of the Fall Symposium of the American Medical Informatics Association, Washington, DC, Nov 5–10, 1999; 970–974.
33. Quinlan J. Learning logical definitions from relations. Machine Learning 1990; 5:239–266.
34. Soderland S, Fisher D, Aseltine J, Lehnert W. CRYSTAL: inducing a conceptual dictionary. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence 1995; 1314–1321.
35. Berger AL, Della Pietra SA, Della Pietra VJ. A maximum entropy approach to natural language processing. Computational Linguistics 1996; 22:39–71.
36. Friedman C, Hripcsak G. Evaluating natural language processors in the clinical domain. Methods Inf Med 1998; 37:334–344.
37. Friedman C, Hripcsak G, Shagina L, Liu H. Representing information in patient reports using natural language processing and extensible markup language. J Am Med Inform Assoc 1999; 6:76–87.
38. Kahn CE Jr. Standard generalized markup language for self-defining structured reports. Int J Med Inf 1999; 53:203–211.
39. Wang C, Kahn CE Jr. Potential use of extensible markup language for radiology reporting: a tutorial. RadioGraphics 2000; 20:287–293.