A Text Visualization Process for the Recognition of Writing Style

Telung Pan[1]
[1]Department of Management and Planning,
St. Joseph's Hospital,
Huwei, Taiwan

Yu-Fang Su[2]
[2]Infection Control Team,
St. Joseph's Hospital,
Huwei, Taiwan

Li-Chin Lu[3]
[3]Medical Research Office,
St. Joseph's Hospital,
Huwei, Taiwan

***Abstract*—This study aims to compare the structure of medical reports of same type medical examination but composed by different physicians. An analysis and compare method was proposed based on formal concept to produce different concept structure for medical reports. Important concepts of medical reports can be described in a hierarchy structure. The concepts and its relational structures can be observed by different concept lattice produced from different physicians. Synonym terms and writing style variations for different concept lattice will be noted from the variations of different hierarchical level and thus can be used for the improve for better medical report writing.**

Keywords- medical report quality; formal concept; synonym.

I. INTRODUCTION

Medical reports of different examination results and interpretations from clinical physicians are important evidences for the correctness of diagnosis. The medical reporting service and quality comes from these important examination reports. Interventions such as computer-added results and standardized format of reports may facilitate more timely and correctness transfer of pertinent patient information to the physician of patients and more correctly during follow-up care [1].

From a medical perspective, some studies reviewed different situations needed to improve the quality of medical reports. In a retrospective study, their analysis of 752 consecutive medical reports of patients with insulin or non-insulin dependent diabetes mellitus, standardized methods of documentations are required for many diabetic inpatients [2]. Another empirical evaluation of the completeness of medical records from 607 randomly selected patients, reveal that the history and physical examination usual lack of key medical elements. There are study shows that patients treated with better hospital medical records quality have significantly lower mortality [3]. The relationship between better medical records and better medical care could be better to monitor patients and improve the quality of medical care.

This study proposed that medical reports is the knowledge and judgment from medical physician's thinking network model, and can be presented in different style of formal concepts. [4].

This study aims to build concept structures from medical report content for same type medical examination but composed by different physicians. It strives to identify the variations for the same type of medical examination report composed by different physicians. These variations found by this study will provide valuable suggestions for the improvement of medical report writing quality.

II. WRITING STYLE AS CONCEPT REPRESENTATION

Medical reports are the only means for the communication between medical examinations and referring physicians. It serves as legal document, as well as a permanent part of the patient's record. The issues for the improvement of the quality of medical report is important for the overall quality of a hospital medical service. In this section, we try to define the good quality of medical report and its structure and lexicon.

A. Medical Report Writing and Defining a Good Report

Despite the importance placed on medical reporting, not much guidelines exist for the proper criteria of good or bad medical reports. The American College of Radiology (ACR) Standard for Communication [5] provides only brief, common sense guidelines for the wording of reports for radiology reports. There is no strict guidance offered to radiologists in the construction of the radiology report begins in medical school training. Minimal formal instruction is provided produced of the lack of consensus on what constitutes a good quality of radiology report [6]. According to studies, the criterions of a good radiology report can be summarized as the Six Cs. [7-9]. Medical reports should be clear, correct, concise, complete, consistent, distinct, and have high confidence level.

The one of most frequent complaint from referring clinic physicians about the understanding of radiology report is poor content and structure. A large analysis of 8,426 reports on chest radiographs found that there are 14 different terms used to describe a single common abnormality, and 23 synonyms for reporting the presence of the same pathology [10].

B. Structure and Lexicon

For more than hundred years, the structures, format, and the content of the medical report have remained unchanged, despite dramatic advances in medical techniques. However, as the progress of computer science, interest is growing in redefining the composition and structure of the medical

© 2015. The authors - Published by Atlantis Press

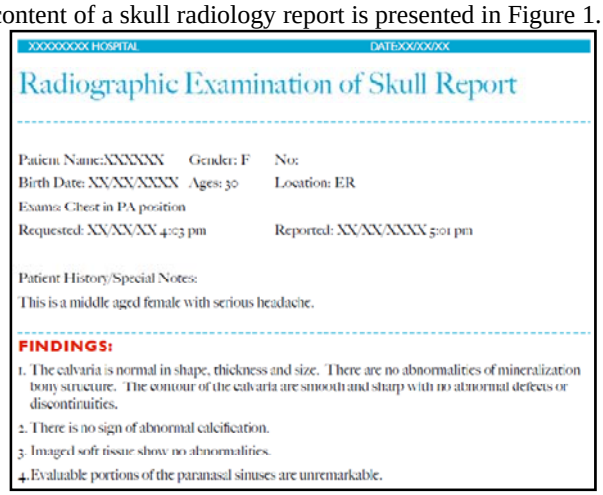
report. With the popular of PACS (Picture Archiving and Communication System), HIS/RIS (Health Information System / Radiology Information System), and the eventual replacement of paper medical records into EMR (Electronic Medical Record), digital tools are now available to redefine the manner and methods in which physicians create medical reports.

In general, any content of formally represented knowledge is based on a conceptualization structure. The objects, concepts and other entities are the elements explicitly or implicitly exist in the knowledge networks that we have [11]. Concept lattice is an explicit specification of a conceptualization of the knowledge networks that we just mentioned [12]. Concept is the basic unit of human thought and also the basic structure of logic. Concepts and its logical structures presented the outside information systems that we know. In this study, we adopt concept lattices based on mathematical abstraction of concept systems to present the human information style and the characteristics of text content [13]. A number of studies attempted to build concept lattice for knowledge representation and applications [14, 15].

III. Medical Report Content Analysis

A. Setting and Raw Data

Research sample raw data were derived from RIS database (radiology information system), an example of the content of a skull radiology report is presented in Figure 1.


XXXXXXXX HOSPITAL DATE:XX/XX/XX

Radiographic Examination of Skull Report

Patient Name:XXXXXX Gender: F No:
Birth Date: XX/XX/XXXX Ages: 30 Location: ER
Exams: Chest in PA position
Requested: XX/XX/XX 4:03 pm Reported: XX/XX/XXXX 5:01 pm

Patient History/Special Notes:
This is a middle aged female with serious headache.

FINDINGS:
1. The calvaria is normal in shape, thickness and size. There are no abnormalities of mineralization bony structure. The contour of the calvaria are smooth and sharp with no abnormal defects or discontinuities.
2. There is no sign of abnormal calcification.
3. Imaged soft tissue show no abnormalities.
4. Evaluable portions of the paranasal sinuses are unremarkable.

Figure 1. Radiology report content of skull examination

B. Term Parsing

The concepts lattices produced from radiology reports are actually a form of writing style characteristic. Differences between concept lattices can be measured by the comparison of their corresponding lattices hierarchical structure and concepts. Document layout, stop words, and stemming words were eliminated in term parsing process.

C. Conceptual Relationship and Hierarchy Construction

There are 3 different relationships between concepts: independence, intersection, and inheritance [16]. The processes for the establishment of concept relationship and hierarchy of medical reports for the representation of its writing structure are shown as follows:

In every writing report, words that best represent the key concepts of the radiology reports have to abstract from its context. This study define two different sets of 'report set' and 'word set', which a formal context *C=(R, W, I)* consists of two sets *R* and *W* and a relation *I* between *R* and *W*. Elements of *R* (reports) are objects, and elements of *W* (words) are attributes of the 'formal context'. An object *r* is in a relation *I* with an attribute *w*, we can write *rIw* or $(r, w) \in I$ and red it as "the object *r* has the attribute *w*'.

A set $A \subseteq R$ of objects we can define

$$A' := \{w \in W \mid rIw \; for \; all \; r \in A\} \quad (1)$$

Correspondingly, for a set *B* of attributes we define

$$B' := \{r \in R \mid rIw \; for \; all \; w \in B\} \quad (2)$$

A formal concept of the context (R, W, I) is a pair (A, B) with $A \subseteq R, B \subseteq W, A' = B \; and \; B' = A$. A is the extent and *B* is the intent of the concept *(A, B)*. *Γ(R, W, I)* denotes the set of all concepts of the context *(R, W, I)*.

If (A_1, B_1) and (A_2, B_2) are concepts of a context, (A_1, B_1) is called a sub-concept of (A_2, B_2), and can denote $(A_1, B_1) \trianglelefteq (A_2, B_2)$. The relation $\trianglelefteq$ is the expression of hierarchical order of the concepts. The set of all concepts of *(R, W, I)* is defined as *Γ(R, W, I)*, which is the concept lattice of the context *(R, W, I)* [17].

D. Many-valued Contexts

In standard language the word 'attribute' is not only used for properties which an object may or may not have. Attributes such as 'color', 'weight', 'sex', and 'height' have measurement degree. We call it many-valued attributes, and is the expand and variant of one-valued attributes [17].

A 'many-valued context' *(R, W, M, I)* has sets of *R*, *W*, and *M*, and a ternary relation I between *G*, *M*, and *W* for which it holds that

$$(r, w, m) \in I \; and \; (r, w, v) \in I \; always \; imply \; m = v \quad (3)$$

E. Interrelation between concepts

After the construction of the hierarchy between concepts, the next process is the definition of interrelation degree between concepts. A concept appear more often in a report document is definitely more important than singly appearing concept in a single report. We can define $r(C_j, C_k)$ as the relationship degree of two concepts.

After the process consisting of the above methods, a matrix table denoting the relation of concepts that appear in medical reports can be completed as the format of formal context, as in Figure 2, in which the conceptual hierarchies generated. Table III is the example of relation degree values between concepts computed from $r(C_j, C_k)$. Combine with the results of

Figure 2 and Table III, an concept lattice with relation degree can be produced as shown in Figure 3.

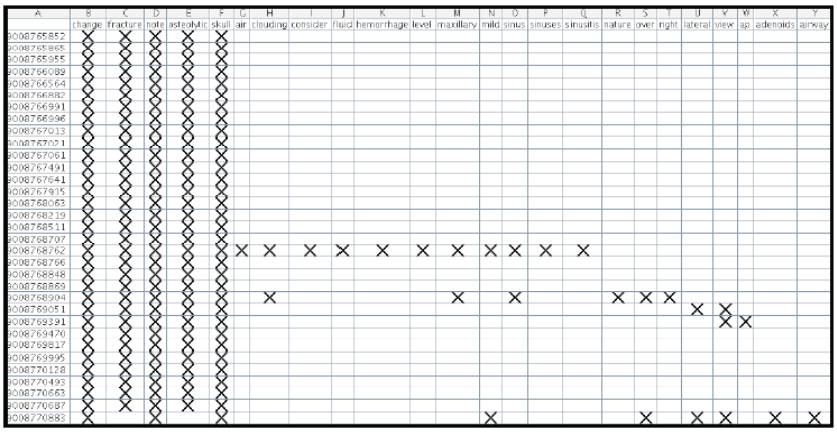


Figure 2. Formal context produced from skull examination reports

The nodes appear in the upper part (root of the trees) of the concept lattice of Figure 3 are the most common concept in different radiology reports, while the nodes appear in the lower part which are more sensitive and variant in different radiology reports. In the concept lattice produced from this study, the hierarchy structure of concepts represent the different nature of writing style of different radiologist or physicians. The graphical structure of concept lattice produced from medical reports composed by different radiologist or physicians can be a useful visual tools for further writing style or quality comparison analysis. In next section, we will show some analysis examples from different medical reports authors.

IV. Writing Style and Synonyms Analysis

In order to evaluate different writing style and detect synonym for different physicians, this study measures the similarity between two given lattice. Different nodes in the concept lattice graph can be treated as different concepts. In this way, we can measure the similarity between different concept lattice charts by the method of weighted graphic matching.

TABLE I. EXAMPLES OF RELATION DEGREE BETWEEN TWO CONCEPTS

ATTR_NAME_A	ATTR_NAME_B	R_VALUE
adenoids	bone	1.06112537370499
adenoids	change	0.71137057952477
adenoids	enlargement	1.06112537370499
adenoids	gross	1.06112537370499
adenoids	lateral	0.99626368868942
adenoids	lesions	1.06112537370499
adenoids	mild	0.99626368868942
adenoids	note	0.64762957414716
adenoids	obstructing	1.06112537370499
adenoids	over	0.99626368868942
adenoids	parietal	1.06112537370499
adenoids	posterior	1.06112537370499

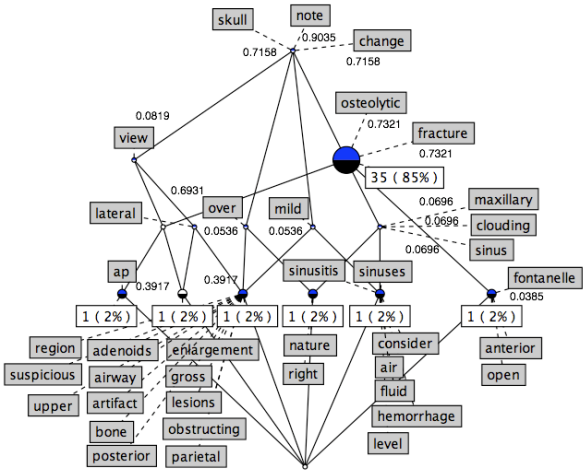


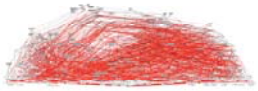
Figure 3. Concept lattice with relation degree

This study compares different skull examination concept lattice for different physicians and finds the variations for different concept lattices. Table II is the result of the measure of similarities, value larger than the others represent more different between two concept lattices. It shows the writing style of radiology RX03 is more different with others. Table III shows the concept lattice charts and the analysis of the writing style and also the synonym observed.

TABLE II. SIMILARITIES FOR DIFFERENT RADIOLOGISTS

Radiologist ID	RX01	RX02	RX03
RX01	0	461	1669
RX02	461	0	2710
RX03	1669	2710	0

TABLE III. THE CONCEPT LATTICE CHARTS AND THE ANALYSIS RESULT OF RADIOLOGY REPORT

RX01	 RX02	RX03
Concepts count: 175 Objects count: 223 Attributes count: 60	Concepts count: 1,108 Objects count: 259 Attributes count: 76	Concepts count: 95 Objects count: 345 Attributes count: 83
Writing style		
1. The description is medium between other radiologist. 2. Fewer same words appear in different reports.	1. The description is simple and the words are fewer than other radiologist. 2. Fewer same words appear in different reports.	1. The description is more complete and detailed. 2. More same words appear in different reports.
Synonym analysis		
H1: [skull]. H2: [change], [note]. H3: [fracture], [over]. H4: [bone],[osteolytic], [right] H5: [left], [metallic] H6: [view],[lateral],[post],[sutures]	H1: [osteolytic, show, sella, skull, lat, change]. H2: [over], [dislocation], [fracture], [ap], [correlate, pls], [clinical],[apparent]. H3: [seen], [some], [soft], [note]	H1: [show, skull, sinus, paranasal, portion, radiographic, film, soft, imp, sign, tissue, exam, image, unremarkable, abnoraml, calcification, evaluable]. H2: [study], [smooth, normal, thickeness, defects, sharp, calvaria, contour, shape], [abnormalities]. H3: [bony, structure, abormality]

V. CONCLUSION

This study proposed a process to analyze medical report content for different physicians of same examination item. The important writing variations of medical report for different physicians are elucidated through the observation of concept lattices. In the level of synonym, the different styles of concept lattices indicates an important point of view that the radiologists can improve the report consistency by using consistent words in writing medical report documents. Our findings suggest that writing style and the usage of words should be systematically and consistently re-organize for different composers to improve the well communication between radiologists and clinical physicians.

REFERENCES

[1] S. Kripalani, F. LeFevre, C.O. Phillips, M.V. Williams, & P. Bassviah, Deficits in Communication and Information Transfer Between Hospital-Based and Primary Care Physicians: Implications for Patient Safety and Continuity of Care, *JAMA*, Vol. 297, No. 8, 2007, pp. 831-841.

[2] B. Liesenfeld, H. Heekeren, G. Schade, & K.D. Hepp, Quality of Documentation in Medical Reports of Diabetic Patients, *International Journal for Quality in Health Care*, Vol. 8, No. 8, 1996, pp. 537-542.

[3] S.M. Dunlay, K.P. Alexander, C. Melloni, J.L. Kraschnewski, L. Liang, W.B. Gibler, M.T. Roe, E.M. Ohman, & E.D. Peterson, Medical Records and Quality of Care in Acute Coronary Syndrome Results from CAUSADE, *Archives of Internal Medicine*, Vol. 168, No. 15, 2008, pp. 1692-1698.

[4] T. Pan, & K. Fang, Ontology-Based Formal Concept Differences Analysis in Radiology Report Impact by the Adoption of PACS, *Lecture Notes in Computer Science*, Vol. 5548, 2009, pp. 267-283.

[5] D.C. Kushner, & L.L. Lucey, Diagnostic Radiology Reporting and Communication: The ACR Guideline, *American College of Radiology*, Vol. 2, No. 1, 2005, pp. 15-21.

[6] K.H. Vydareny, Are Today's Residents Ready for Tomorrow's Practice? *American Journal of Roentgenology*, Vol. 173, No. 3, 1999, pp. 537-538.

[7] R.R. Armas, Qualities of a Good Radiology Report, *American Journal of Roentgenology*, Vol. 169, No. 4, 1997, pp. 943-946.

[8] M. Lafortune, G. Breton, & J.L. Baudouin, The Radiological Report: What is Useful for the Referring Physician? *Canadian Association of Radiologists Journal*, Vol. 39, No. 2, 1988, pp. 140-143.

[9] R.F. McLoughlin, C.B. McLoughlin, C.B. So, R.R. Gray, & R. Brandt, Radiology Reports: How Much Descriptive Detail is Enough? *American Journal of Roentgenology*, Vol. 165, 1995, pp. 803-806.

[10] J.L. Sobel, M.L. Pearson, K. Gross, K.A. Desmond, E.R. Harrison, L.V. Rubenstein, W.H. Rogers, & K.L. Kahn, Information Content and Clarity of Radiologists' Reports for Chest Radiography, *American Journal of Kidney Diseases*, Vol. 3, No. 9, 1996, pp. 709-717.

[11] M.R. Genesereth, & N.J. Nilsson, *Logical Foundations of Artifical Intelligence*, 1987, San Mateio: Morgan Kaufman Publishers.

[12] R. Wille, Formal Concept Analysis as Mathematical Theory of Concepts and Concept Hierarchies, *Formal Concept Analysis*, *LNCS*, Vol. 3626, 2005, pp. 1-33.

[13] R. Wille, Why Can Concept Lattices Support Knowledge Discovery in Database? *Journal of Experiment & Theoretical Artificial Intelligence*, Vol. 12, No. 2, 2002, pp. 81-92.
[14] B. Ganter, Attribute Exploration with Background Knowledge, *Theoretical Computer Science*, Vol. 217, No. 6, 1999, pp. 215-233.
[15] T. Pan, & K. Fang, Ontology-based Formal Concept Differences Analysis in Radiology Report Impact by the Adoption of PACS, *Formal Concept Analysis, LNCS*, Vol. 5548, 2009, pp. 267-283.
[16] B. Ganter, G. Stumme, & R. Wille, *Formal Concept Analysis: Foundations and Applications*, 2005, Heidelberg: Springer.
[17] B. Ganter, & R. Wille, *Formal Concept Analysis: Mathematical Foundations*, 1999, Heidelberg: Springer.