Automatic generation of repeated patient information for tailoring clinical notes

Frank Meng[a,*], Ricky K. Taira[a], Alex A.T. Bui[a],
Hooshang Kangarloo[a], Bernard M. Churchill[b]

[a] *Medical Informatics Group, University of California, Los Angeles, 924 Westwood Blvd.,
Suite 420, Los Angeles, CA 90024, USA*
[b] *Department of Urology, University of California, Los Angeles, Los Angeles, CA 90024, USA*

Summary   Generating clear, readable, and accurate reports can be a time-consuming task for physicians. Clinical notes, which document patient encounters, often contain a certain set of patient information including demographics, medical history, surgical history, examination results or the current medical condition that is propagated from one clinical note to all subsequent clinical notes for the same patient. To this end, we present a system, which automatically generates this patient information for the creation of a new clinical note. We use semantic patterns and an approximate sequence matching algorithm for capturing the discourse role of sentences, which we show to be a useful feature for determining whether the sentence should be repeated. Our system is shown to perform better than a simple baseline metric using precision/recall results. We believe such a system would allow clinical notes to be more complete, timely, and accurate.

1. Introduction

Within a series of reports for a given patient, a percentage of the patient's information is often repeated, being carried over from one report in the series to the next. Physicians often spend much time and effort both determining what information needs to be repeated and re-generating (dictating or typing) this repeated information when creating a follow-up report for a given patient [1]. This paper describes a methodology to automatically generate repeated patient information for creating new clinical notes. We define a clinical note to be a report which documents encounters with patients, including information such as demographics, medical history, surgical history, examination results or the current medical condition. Our belief is that such a system can reduce the total amount of time needed to generate clinical notes and can also lead to more complete and accurate notes. Completeness is achieved because the system will always propagate to subsequent notes information designated by the physician to be repeated. The accu-

* Corresponding author.
   E-mail address: fmeng@mii.ucla.edu (F. Meng).

racy of the repeated information is dependent on the accuracy of the previous note(s) in the series but there is the ever-present danger of propagating erroneous information. With a user interface, which clearly presents the repeated information to the physician, the content of each clinical note can be thoroughly reviewed before being submitted to the patient's permanent record.

Our approach is to use the role a phrase or sentence plays within the document as a key feature in determining whether it should be repeated or not. The role a sentence plays, called a discourse role, is basically the author's intention in using that sentence. We have identified several discourse roles that occur within clinical notes in the pediatric urology domain and some were found to be repeated more readily than others. The discourse role and other words mentioned within the same sentence were shown to be good features for predicting repetition of sentences in clinical notes with high accuracy. Though we have designed the system with our particular clinical note in view, we believe that our methodology is applicable to other documents, which share similar characteristics.

2. Methods

2.1. Generating repeated patient information

Our system generates text for documents by extracting text segments from a previous document and inserting it into a new document. We believe copying text verbatim for generating repeated patient information is suitable in performing this task for the following two reasons:

(1) Physicians often use particular phrases to describe medical observations or events and a system, which generates its own language, may incorrectly convey this vital information.
(2) There is usually no need for the system to generate new text because in practice, repeated information is often propagated over verbatim into a new document by the physician.

Fig. 1 shows the overall functional diagram of the system we describe in this paper. The system requires a past clinical note as input and, using the information contained in it, generates the repeated patient information for the subsequent note. Repeated patient information can contain static patient information such as the patient's name, age, and known medical history, or it can contain information which may change from patient to patient, such as the patient's primary symptoms, current ill-
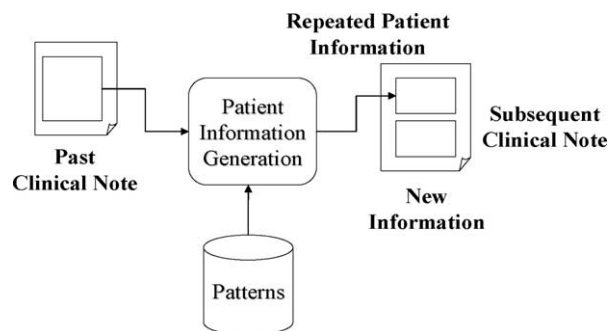

Fig. 1 System diagram.

nesses, and the patient's treatment plan. Repeating static information is a rather trivial task, which can be accomplished by standard structured reporting techniques such as templates or macros. However, information that differs from patient to patient requires that the system have a deeper understanding of the language contained in the medical document. The following is an excerpt from a typical clinical note authored by a pediatric urologist:

This patient has very profound general problems. He was born with the cord doubly wrapped around his neck just before birth. There was severe loss of fetal heart rate. He was delivered by cesarean section, but unfortunately has been severely retarded and quadriplegic since then. He is always in a wheelchair. He is not able to feed or dress himself. He does not respond in any meaningful way. He has no language skills. He has had extensive trouble with gastroesophageal reflux and is now fed entirely by a gastrostomy tube. He also has had severe scoliosis for which there has been an anterior fusion. More recently, the problem has been one of hematuria and blood in the urine.

Typically, a clinical note begins by indicating the major presentations of the patient. As the note progresses, new information is added, such as the purpose of the current visit and any new findings. The following is an excerpt from a clinical note authored by the same urologist after the next encounter with the same patient:

This patient had a severe problem at birth with the cord double wrapped around his neck when he was born. There was severe loss of fetal heart rate and he suffered irreversible ischemic cerebral damage. He has significant problems with severe retardation. He is quadriplegic. *He is fed with a gastrostomy tube.* We have been having a number of urinary and fecal problems with this patient.

The italicized and bolded portions of the subsequent note above show the patient information that was carried over from the previous note, which is a