Performing audit of manual EMR coding results

* Background
  + Medical records
  + Important information but not readily digestible by computer sytems
  + Thus manual coding industry
* Challenges in NLP of medical records
  + Most general alrogithms don’t valid (more for email classification or relevance ranking)
  + Negation
  + Cost of error
* Automated systems accuracy vs manual
  + General study
  + Case 1
  + Case 2 Regnestrief
  + Case 3 MP – sepsis
    - Example 1
    - Example 2
    - Overall metrics
* Conclusion and Outlook

Understanding of Electronic Medical Records(EMRs) plays a crucial role in improving healthcare outcomes. However, the unstructured nature of EMRs poses several technical challenges for structured information extraction from clinical notes leading to automatic analysis. Natural Language Processing(NLP) techniques developed to process EMRs are effective for variety of tasks, they often fail to preserve the semantics of original information expressed in EMRs, particularly in complex scenarios. This paper illustrates the complexity of the problems involved and deals with conflicts created due to the shortcomings of NLP techniques and demonstrates where domain specific knowledge bases can come to rescue in resolving conflicts that can significantly improve the semantic annotation and structured information extraction. We discuss various insights gained from our study on real world dataset.

<http://dl.acm.org/citation.cfm?id=2512427>

## Background

* 02-NLP is the formulation and investigation of comuptatinonlly effective mechanisms for communication through natuarla language
* 02-Named entity recognition is a sub-field of inromation extraction and referes to the task of reconigzing expressions denoting entities (named entites) such as diseases, druges, or people’s names, in free text docuemts
* 02-Ruled-based NER systems can be very effective, vut require some manual effort. ML appreoaches can successfully extract named entities but require large annotated training corpora. Advnatages of ML approaches are that they do not require human inutuition and can be retrained wihotut reporgramming for any domain
* 02-Text mining uses information extraction is is defined by Heasr t as the process of discovering and extracting knowledge from unstrcutred data. Text mining typically ocmpresis two or three steps: information retrial (to gather relevant etexxts); the step is not alwaysnecessary), information extraction ( to extract specfici typies of information from texts of interst), and data mining (to find associations among the extracted pieces of information)
* 01-Clinical coding and classification processes transform natural language descritpions in clinical text into data that can subsequently be used for clinical care, research, and other purposes
* 01-Automated coding and classification systems themselves are not generalizable, nor are the results of the studies evaluating them.
* 01-Automated coding and classification technologies encompass a variety of computer-based approaches that transform narrative text in clinical records into structured text, which may including assignement of codes from standard terminologies, without human interaction.
* 01-not clear whether these automated systems perfrom as well as manual coding or classification
* 01 (2009)-there are no systematic reviews on automated clinical coding and classification systems. Meystre et al conducted a narrative review to examine published research on the extraction of information from textual documents in the electoric health records. In that review NLP techniques were EXAMINED, BUT FEW OF the studes dealt specifically with automate coding and classification software
* 01-[this paper] is a systematic literature review to idneitfy all published studies evaluating the performance of automated coding and cliassificaiton systems
* 01-The use of structured data in coded form continues to grow as the healthcare industry explores value-based purchasing and seek overall improvement in the quality of care. The data used for thse purposes are typically encoded via a mnaul coding process. This process involves human review of clinical documentation to identify applible codes..[usually be coding professionals]
* 01-‘The industry needs automated solutions to allow the coding process to become more productive, efficient, accurate, and consistent” according to AHIMA
* 01-computer applications for automating this process are available but currently not widely used most likely because the systems are still in development and their performance in production unproven
* 01-Finds 46 different systems named in review of 113 academic studies indexed in Medline or other relevant databases prior to March 2009. Most common was Columbia University’s MedLEE, followed by SymText, MMTx, and NegEx. These four systems toegether represtn 91% of the named systems studied and 37% of the total corpus
* 01-three reference standards to judge automated systems:
  + Gold standar—2 or more independent reviewers with ajdjudication of diagreements
  + Trained standard: one epxert
  + Regular practice: on human reviewer
* 01-broad range of coding classification systems
* 01-Of the 113 studies included in our review, 26 specifically asserted tha the automated system performed better than, or as well as humans, while only four explicitly stated that humans outperformed the automated systems

## Potential

* 02-Growth of narrative data in electronic form, along with the needs for improved quality of care and reduced medical errors are both strong incentives for the development of NLP
* 02-Much of the available clinical data are in narrative form as a result of transcirpiton of dictations, direct entry by providers, or use of speech recognition applications. This free-text form is convenient to express concepts and events but is diffulct for searching ,summarization, decision-support, or staitictical analsysi
* 02-to redue reerors and improve quality control, coded data are reuired; this is where NLP, and more precisely information extraction is neded
* 02-information extracted can then be linked to concepts in standard terminologies and cused for coding

## Challenges

* 01-2006, Kukafka observed that ‘coding tasks involving complex reasoning such as those in which disparate pieces of information must be connected, are a difficult challenge for current NLP systems
* 01-medical NLP tools…re difficult to adapt, generalize, and re-use…’a new set of regular expressions has to be developed and validated for each particular task’
* 01- Barrows et al states, ‘as if NLU of narrative text docuemtns by computer systems is not difficult enough, the understanding of notational text docuemtns is perhaps even more difficult due to lack of punctuation and grammear and frequesnt use of terse abbrreviations and symbols
* 02-since narratives are rarely made available outside the corporate setting that generated them, formal studies of them are sparase
* 02-some clinical texts are ungrammatical and comprosed of short, telegraphic phrases
* 02-Narratives are pregnant with short-hand (abbreviations, acronyms, and local dialectal shortahand prhases). These shorthand lexical units are often overloaded (est 33%)
* 02-misspellings
* 02-clinical narratives coan contain any charatcters that can be typed or pasted
* 02-explicit tempaltes are pre-formatted, highly idiosyncratic, and instituiotn-specific with fields to be filled in by the user

## Process

* **02 is great resource for this**
* 02-IE typically reuires some pre-pocessing such as spell checking, document structure analysis, sentence splitting, tokenization, word sense disambiguation, part of speech tagging, and some form of parsign
* 02-typical process consists of combination of the following compoenents:
  + Tokenizer
  + Sentence boundary detector
  + Art-of-speech tagger
  + Morphological analyzer
  + Shallow parser
  + Deeper parser (optional)
  + Gazetteer
  + Named entity recognizer
  + Discource module
  + Template extractor
  + Template combiner

## **Conculsions**

* 01-we conclude from this systematic literature review that automated clinical coding and classification system performance is relative to the complexity of the task and the desired outcome. Automated coding and classification systems themselves are not generalizable and neither are the evaluation results in the studies. More work to correleate tehe purpose and relate complexity of these studies with evaluation results could be informative as weould further analysis to determineif performance of automated systems has remained static over time or if the lack of obvious statitcal improvement is a reflection of more and more diffiuclt tasks being attempted by the automated sytems under evaluation
* The published research examined this review shows that automated coding and cliassfication systems hold promise but the application of automed coding must be considered in context

A Field Theoretical Approach to Medical Natural

Language Processing

* Preliminary performance, as quantified by link recall and precision statistics, is 84.9% and 89.9%, respectively
* The goal of the medical natuarla language processing (NLP) is to transform the information content contained within a free text report (e.g., radiolog) into a representation that is computer understable.
* Several high-end applications have been developed for a variety of high-end aplications including automatic coding of patient reports [3], [4], extraction of findings documented in diagnostic reports [5]-[8], automatic flagging of alarm conditions [10], and analysis of co-occurrence relations among radiological findings [11].

A Novel System for the Automatic Extraction of aPatient Problem Summary

* Considerin that medical information comes from multiple sources, a system fo the automatic generation of pboelm lists could prove to be very effective in terms of saving time in the analysis of large amounts of medical data
* Natural Language Processing (NLP) can support the process of data organization and automated problem list generation, unlocking the precious information embedded in clinical narratives. Information extraction is a common application of NLP in biomedicine, consisting in the identification of relevant entities, toegether with the interpretation of modifiers and relatinos between them [2]. NLP techniques can be used to extract clinical named entities and automatically generate patient pboelm lists for care providers.
* A system able to provide the clinician with a rapid snapshot of the patient’s salient information, reporting a prolem list, with references to mmedications taken, and to the procedures employed for patient management. Moreover, the system includes the assignment of standard codes to recognized disorder [4]. In particular disorders are codified according to the International Classifiation of Diseases version 9 (ICD-9).
* In medicine, Electronic Health Record (HER) systems collect important clinical events in a digital formt
* The first component of the Extraction Subsystem is the NLP pipline,used to extract relevant information from narrative text…in particular the NLP pipline includes the following stages:
  + Language identification
  + Tokenization
  + Lexical analysis
  + Parsing rules
  + Medical entity recognition
* A practitioner generalyy has available a large number of docuemnts regarding the condition of patients under treatment. Analyzing the entirety of such records is time-consuming. A system, able to automatically summarize the most important clinical information relating to a given patient, can help physicians to obtain easily a clnical picture of his/her health status, contributing to an efficient process of care.

**A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries**

* Narrative reports in medical records contain a wealth of information that may augment structured data for managing patient information and predicting trends in diseases
* Pertinent negatives are evident in text but are not usually indexed in structured databases
* We developed a simple regular expression algortithm called NegEx that implements several phrases indicating negation, filters our sentences containing phrases that falsely appear to be negation phrases, and lmits the cope of the negation phrases.
* Much of the clinical information contained in patient medical reocrds is in narrative form and therefore unabailable to automated systems that could improve patient care or further medical research
* Infomration retrieval techniques, however, do not generally discriminate between terms that are mentioned as being present and terms thata re negated
* Researchers in the medical language p rocessing community have create methods for automatically extracting information contained in narrative reports for diceions support [5], guidelines implementation [6,7], detection and management of epdicemics [8], and identification of patients eligible for research studies [9].

**Use of General-purpose Negation Detection to Augment Concept Indexing of Medical Documents**

* **Results:** In the first evaluation using marked-up documents, 8,358 instances of UMLS concepts were

detected in the 60 documents, of which 544 were negations detected by the program and verified by

human observation (true-positive results, or TPs). Thirteen instances were wrongly flagged as negated

(false-positive results, or FPs), and the program missed 27 instances of negation (false-negative results,

or FNs), yielding a sensitivity of 95.3 percent and a specificity of 97.7 percent. In the second evaluation

using independent negation detection, 1,869 concepts were detected in 10 documents, with 135 TPs, 12

FPs, and 6 FNs, yielding a sensitivity of 95.7 percent and a specificity of 91.8 percent. One of the words

“no,” “denies/denied,” “not,” or “without” was present in 92.5 percent of all negations.

* Conclusions: Negation of most concepts in medical narrative can be reliably detected by a simple

strategy. The reliability of detection depends on several factors, the most important being the

accuracy of concept matching.

* Automated concept indexing based on the NationalLibrary of Medicine’s Unified Medical Language System (UMLS) Metathesaurus3 or its MeSH (Medical Subject Headings) subset has been explored by several researchers.4–8
* An important aspect of information-retrieval-based search is the ranking of matching documents by relevance, 9,10 giving more weight to documents containing the specified keywords many times and to documents that contain keywords that are rarer in the collection as a whole. For a medical document, however, the presence of a concept does not necessarily make the document relevant for that concept. The concept may refer to a finding that was looked for but found to

be absent or that occurred in the remote past.

* Therefore, to increase the utility of concept indexing of medical documents, it is necessary to record whether the concepts have been negated or not.

Manual versus automated coding of free-text self-reported medication data in the 45 and Up Study: a validation study

* **Methods:** A random sample of 500 participants (475 with and 25 without medication data in the free-text box) enrolled in the 45 and Up Study was selected. Manual coding involved medication experts keying in free-text responses and coding using Anatomical Therapeutic Chemical (ATC) codes (i.e. chemical substance 7-digit level; chemical subgroup 5-digit; pharmacological subgroup 4-digit; therapeutic subgroup 3-digit). Using keyed-in free-text responses entered by non-experts, the automated approach coded entries using the Australian Medicines Terminology database and assigned corresponding ATC codes.
* **Results:** Based on manual coding, 1377 free-text entries were recorded and, of these, 1282 medications were coded to ATCs manually. The sensitivity of automated coding compared with manual coding was 79% (*n* = 1014) for entries coded at the exact ATC level, and 81.6% (*n* = 1046), 83.0% (*n* = 1064) and 83.8% (*n* = 1074) at the 5, 4 and 3-digit ATC levels, respectively. The sensitivity of automated coding for blank responses was 100% compared with manual coding. Sensitivity of automated coding was highest for prescription medications and lowest for vitamins and supplements, compared with the manual approach. Positive predictive values for automated coding were above 95% for 34 of the 38 individual prescription medications examined.
* **Conclusions:** Automated coding for free-text prescription medication data shows very high to excellent sensitivity and positive predictive values, indicating that automated methods can potentially be useful for large-scale, medication-related research.