Chi-Ming Chu
Dr. sc. hum.

Case-Based Reasoning Adverse Drug Events Detection with Medical Language Processing and Artificial Neural Iteration Clustering

Geboren am 15.05.1970 in Kaohsiung
Reifeprüfung am 01.07.1988 in St. Dominic High School Kaohsiung
Studiengang der Fachrichtung Öffentliches Gesundheitswesen vom 1988 bis 1994
Vordiplom am 01.07.1992 an der Universität, Kaohsiung Medical College
Diplom am 01.07.1994 and der Universität, National Defense Medical College
Promotionsfach: Medizinische Informatik
Doktorvater: Prof. Dr. Thomas Wetter

Case-Based Reasoning Event Detection (CBRED) techniques can identify the onset of new events from data or information by scanning narrative documents, which convey information about the respective events. The CBRED techniques can be applied to the medical field, i.e., for the identification of potential events by comparing the similarity of words between newly coming and past narrative documents. CBRED is a major extension of the traditional Feature-Based Inspecting Event Detection (FBIED) techniques that turned out to be severely limited when applied in medical informatics.
Therefore, the thesis will address the inherent problems of FBIED by proposing the CBRED techniques as a combination of information extraction and text categorization techniques to overcome the limitation of the traditional Information Extraction (IE) because an IE machine usually applies its domain knowledge based on events. It is often limited in its application to one specific domain, because the event detection works on the basis of predefined keywords The predefined-term limitation will be overcome in this thesis by introducing a combination of unsupervised learning of pertinent terms, neural network formation of major concepts and intelligent information clustering for allowing to identify known event narratives that closely match a new event description. In the proposed model two kinds of narratives are captured: discharge summaries and drug information documents.
In order to detect the hidden events in documents effectively, Natural Language Processing (NLP) is applied, or more specifically sub-language processing in the form of Medical Language Processing (MLP), which can give support to the selection of medical terms, the connection of synonyms and the semantics of such terms. Before MLP is applied words that are not terms (such as *in, some, after)* are filtered by an analysis based on the Vector Space Model (VSM). What remains are medically meaningful terms. These are selected and mapped onto concepts of the Unified Medical Language System (UMLS) Metathesaurus. UMLS is the major MLP resource used in the thesis. These concepts in the Metathesaurus are supposed to appropriately characterize drug event related contents in our medical narrative documents. Consequently, the thesis also utilizes the multilingual concept identifiers in the Metathesaurus to link documents from the two different languages (German and English).
For giving advice on the basis of thus captured previous drug event narratives, Expert Systems techniques can be useful; more precisely, a Case-Based Expert System, and a Knowledge Base representing known events as cases. However, it has not been clarified how to construct the case base such that the user can effectively retrieve the appropriate case; this has been a known bottleneck in building case-based expert systems. For the respective solution, a knowledge representation is proposed for the description of relationships among all medical narrative text segments, which can be effectively managed by appropriate combination of MLP and

IE. After the case base has been constructed by making use of the paragraphs that constitute the document structure, the hierarchy of all cases is also constructed. Instead of the traditional flat structure of narratives in a case base, the cases' hierarchy and preprocessed clusters help enhancing the efficiency of case retrieval. Based on this ideal of a "structured document", formal case construction, measures for similarity and document clustering with artificial neural network models have been set up to find that case in the collection of clinical narratives and drug information documents that is most similar to the new case.
In principle the above methods can be applied to all kinds of medical events described in medical professional language. Adverse Drug Events (ADEs) - such as Drug-Drug Interactions- are an important and costly problem in medical treatment. Since a timely and effective detection cannot be guaranteed in routine clinical care, it causes that many events are not reported. The thesis presented here proposes approaches for the implementation of a better detection of contraindicated Drug-Drug Interactions which can serve three purposes: (1) clinicians are alerted about the risk of Drug-Drug Interactions entailed by medications they are planning, (2) detailed information concerning Drug-Drug Interactions is given to clinicians as well as to pharmacists, and (3) pharmacists are supported in both checking the prescriptions and providing clinicians with essential information. A document scanning is the essential prerequisite for acquisition and usage of information about events, trends, and relationships among documents. It can offer information to detect documental problems that will improve users' awareness to correct possible errors and monitor organizational service. For example, the documental scanning of clinical notes can alert clinicians' awareness of potential medications with Drug-Drug Interactions to adjust, improve and monitor the quality of patients' care.

Keywords: Narrative text, information extraction, information retrieval, event detection, natural language processing, medical language processing, case-based reasoning, neural network clustering