Chapter XI
Analyzing the Text of Clinical Literature for Question Answering

Yun Niu
Ontario Cancer Institute, Canada

Graeme Hirst
University of Toronto, Canada

ABSTRACT

The task of question answering (QA) is to find an accurate and precise answer to a natural language question in some predefined text. Most existing QA systems handle fact-based questions that usually take named entities as the answers. In this chapter, the authors take clinical QA as an example to deal with more complex information needs. They propose an approach using Semantic class analysis as the organizing principle to answer clinical questions. They investigate three Semantic classes that correspond to roles in the commonly accepted PICO format of describing clinical scenarios. The three Semantic classes are: the description of the patient (or the problem), the intervention used to treat the problem, and the clinical outcome. The authors focus on automatic analysis of two important properties of the Semantic classes.

INTRODUCTION

The vast increase in online information brings new challenges to the area of information retrieval (IR) in both query processing and answer processing. To free the user from constructing a complicated boolean keywords query, a system should be able to process queries represented in natural language. Instead of responding with some documents relevant to the query, the system should actually answer the questions accurately and concisely. Systems with such characteristics

are *question-answering* (QA) systems, which take advantage of high-quality natural language processing and mature technologies in IR. The task of a QA system is to find the answer to a particular natural language question in some predefined text. In this paper, we propose an approach that aims to automatically find answers to clinical questions.

Clinicians often need to consult literature on the latest information in patient care, such as side effects of a medication, symptoms of a disease, or time constraints in the use of a medication. The published medical literature is an important source to help clinicians make decisions in patient treatment (Sackett & Straus, 1998; Straus & Sackett, 1999). For example:

• **Q:** In a patient with a suspected MI does thrombolysis decrease the risk of death if it is administered 10 hours after the onset of chest pain?

An answer to the question can be found in *Clinical Evidence* (CE) (Barton, 2002), a regularly updated publication that reviews and consolidates experimental results for clinical problems:

• **A:** Systematic reviews of RCTs have found that prompt thrombolytic treatment (within 6

hours and perhaps up to 12 hours and longer after the onset of symptoms) reduces mortality in people with AMI and ST elevation or bundle branch block on their presenting ECG.

Studies have shown that searching the literature can help clinicians answer questions regarding patient treatment (Cimino, 1996; Gorman, Ash, & Wykoff, 1994; Mendonça, Cimino, Johnson, & Seol, 2001). It has also been found that if high-quality evidence is available in this way at the point of care—e.g., the patient's bedside —clinicians will use it in their decision making, and it frequently results in additional or changed decisions (Sackett & Straus, 1998; Straus & Sackett, 1999). The practice of using the current best evidence to help clinicians in making decisions on the treatment of individual patients is called *evidence-based medicine* (EBM).

Clinical questions usually represent complex information needs and cannot be answered using a single word or phrase. For a clinical question, it is often the case that more than one clinical trial with different experimental settings will have been performed. Results of each trial provide some evidence on the problem. To answer such a question, all this evidence needs to be taken into account, as there may be duplicate evidence,

Figure 1. Example of a clinical question, with corresponding evidence from Clinical Evidence

---

Clinical question: Are calcium channel blockers effective in reducing mortality in acute myocardial infarction patients?

Evidence 1: … calcium channel blockers do not reduce mortality, may increase mortality.

Evidence 2: … verapamil versus placebo … had no significant effect on mortality.

Evidence 3: … diltiazem significantly increased death or reinfarction.

Evidence 4: … investigating the use of calcium channel blockers found a non-significant increase in mortality of about 4% and 6%.

Related Content

A Framework for Multidimensional Real-Time Data Analysis: A Case Study for the Detection of Apnoea of Prematurity
Christina Catley, Kathy Smith, Carolyn McGregor, Andrew Jamesand J. Mikael Eklund (2013). *Methods, Models, and Computation for Medical Informatics (pp. 16-35).*
www.irma-international.org/chapter/framework-multidimensional-real-time-data/73069

The Humane Dimensions of Effective Communication
Bruce Hugman (2012). *Pharmacoinformatics and Drug Discovery Technologies: Theories and Applications (pp. 24-43).*
www.irma-international.org/chapter/humane-dimensions-effective-communication/64064

A Latent Feature Model Approach to Biclustering
José Caldasand Samuel Kaski (2016). *International Journal of Knowledge Discovery in Bioinformatics (pp. 11-28).*
www.irma-international.org/article/a-latent-feature-model-approach-to-biclustering/172003

About Neural Networking of Avatar-Based Modeling as an Intellectual Tool and Technology for Biomedical Study in Bioinformatics
Vsevolod Chernyshenko (2019). *International Journal of Applied Research in Bioinformatics (pp. 57-63).*
www.irma-international.org/article/about-neural-networking-of-avatar-based-modeling-as-an-intellectual-tool-and-technology-for-biomedical-study-in-bioinformatics/237202

A Review of Kernel Methods Based Approaches to Classification and Clustering of Sequential Patterns: Sequences of Discrete Symbols
T. Veena, A. D. Dileepand C. Chandra Sekhar (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications (pp. 72-92).*
www.irma-international.org/chapter/review-kernel-methods-based-approaches/76057