Computers in Biology and Medicine 73 (2016) 71–93

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Data-mining to build a knowledge representation store for clinical decision support. Studies on curation and validation based on machine performance in multiple choice medical licensing examinations

Barry Robson [a,b,c,*,1], Srinidhi Boray [a,b]

[a] Ingine Inc., DE, USA
[b] The Dirac Foundation clg, Oxfordshire, UK
[c] St. Matthew's University School of Medicine, Cayman Islands

ARTICLE INFO

Article history:
Received 18 November 2015
Received in revised form
5 February 2016
Accepted 17 February 2016

Keywords:
Clinical decision support
Data mining
Knowledge representation
Natural language processing
Web surfing
Probability
Relevance
Medical licensing examinations
USMLE

ABSTRACT

Extracting medical knowledge by structured data mining of many medical records and from unstructured data mining of natural language source text on the Internet will become increasingly important for clinical decision support. Output from these sources can be transformed into large numbers of elements of knowledge in a Knowledge Representation Store (KRS), here using the notation and to some extent the algebraic principles of the Q-UEL Web-based universal exchange and inference language described previously, rooted in Dirac notation from quantum mechanics and linguistic theory. In a KRS, semantic structures or statements about the world of interest to medicine are analogous to natural language sentences seen as formed from noun phrases separated by verbs, prepositions and other descriptions of relationships. A convenient method of testing and better curating these elements of knowledge is by having the computer use them to take the test of a multiple choice medical licensing examination. It is a venture which perhaps tells us almost as much about the reasoning of students and examiners as it does about the requirements for Artificial Intelligence as employed in clinical decision making. It emphasizes the role of context and of contextual probabilities as opposed to the more familiar intrinsic probabilities, and of a preliminary form of logic that we call presyllogistic reasoning.
© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Medical knowledge in computer systems

The growth of the ability of computers to capture and use clinical and biomedical knowledge may represent an important transition in human history [1]. In particular, the wealth of data and knowledge on the Internet and its World Wide Web should lead to improved clinical decision support (CDS) by computer systems, i.e. to improved clinical decision support systems CDSS [1]. Prior to the growth of the Internet, software with similar goals, such as that of the pioneering Stanford MYCIN project [2] did, of course, exist, and it is notable that right from the outset, most such systems developed for medicine were seen as needing to consider probabilistic measures, such as degrees of certainty, to be associated with statements of clinical or biomedical knowledge [1,2]. However, these were *Expert Systems* that obtained their knowledge offline by useful statements about the world inputted with associated probabilities estimated by *human experts*, often seen as requiring a specialist human *knowledge engineer* to act as mediator, and overall representing a very time-consuming process [1,2]. We recently introduced a CDS application called MARPLE [3]. MARPLE stands for *Medical Automated Reasoning Programming Language Environment*. A common theme of work of this kind is that it involves a repository of knowledge in a form that computers can more readily use. Such a repository for prediction and decision making is said to be a *knowledge representation store* (KRS). Any kind of KRS is a set of syntactic and semantic conventions that describe things and relationships. Any specific example from such a store is a knowledge element or *KRS element*, that early Expert System designers might describe as a kind of *frame* [1,2]. MARPLE rests on a considerable body of previous work by ourselves, collaborators, and other workers, and these efforts close to its essential features are first reviewed here (this Section 1). We also introduce the new version, MARPLE 2, which has significant advantages in helping ensure the quality of the above knowledge, an important consideration as follows.

* Corresponding author.
E-mail address: robsonb@aol.com (B. Robson).
[1] Tel.: +1 345 928 7242; fax: +1 345 945 3130.
www.ingine.com, www.diractfoundation.org.

http://dx.doi.org/10.1016/j.compbiomed.2016.02.010
0010-4825/© 2016 Elsevier Ltd. All rights reserved.

1.2. The impact of data mining

Many matters discussed in this paper are not yet widely seen as significant pressing problems for the current CDSS industry, because the types of decision support that are currently most widely used are still largely limited to alerts, reminders, and tools designed to ease workflow or enhance cognition [4]. However, studies like MARPLE are timely. There has been an escalating interest in "Big Data" and rapid progress in *data mining* of it [1], including of electronic health records [1,5]. The input data being "mined" is usually conveniently classified as of two types (a) *structured data* as in spreadsheets but also including *relatively structured data* as in electronic health records (EHRs) or similar public health sources, and (b) so-called *unstructured data*, in the present case simply meaning that it is mainly represented by natural language text (NLT) on web pages and other medical text accessible in digital form. MARPLE gets some of its knowledge to use, test, or further curate if necessary from offline digital repositories of data and other knowledge collections [1,4,5], and from the automatic "surfing" of the Internet, particularly to extract knowledge from NLT as described in this paper. Because of the escalating quantity of information obtained by data mining, and with future applications to CDS in a real clinical setting in mind, the quality of it as usable and authentic knowledge is of concern. MARPLE draws on both structured and unstructured data not least because each has well known strengths and weaknesses. Notably, while structured data mining can efficiently provide probabilities to certain important kinds of knowledge, mining natural language text on the Internet usually faces the problem that prevalence may reflect matters of interest and newsworthiness in inverse relation to actual frequency of occurrences in the "real world". Combining elements of knowledge from various sources and exploring means of overcoming the above kind of probability problem are by no means unique to our efforts (e.g. Ref. [6]). Nonetheless, the escalation of collected knowledge makes it difficult to keep up with ensuring its quality.

1.3. Curation

Considerable focus in this present paper is placed on methods of ensuring good *provenance* of KRS elements, which is essentially a matter of demonstrating and ensuring the above quality. Especially with applications to CDS in real clinical settings in mind, knowledge elements should come as much as possible from good sources and adequately represent the originally intended information, but in a form usable by computers. In our definition, all aspects of this including tidying, correction, and even rejection if necessary, represent *curation*, briefly reviewed in Section 1.8. The MARPLE project is primarily a study to develop better methods for curation of KRS elements for CDS. In addition to structured and unstructured data mining, MARPLE also gets some of its knowledge from human experts just as was the case in Expert Systems like MYCIN [2] and as persists for many CDSS today [1,5]. However, with MARPLE, the role of the human expert has primarily become one of *auditor*, and of *curation not creation* of KRS elements already obtained automatically by data mining. Compared with early Expert Systems, the human role is now more responsive rather than proactive. This is important because having human experts provide knowledge in good form is long known to be time-consuming [1,2], and it is easier to automate a task of this well-defined nature.

1.4. Using medical licensing examinations

The prominent, unusual, and perhaps controversial feature of MARPLE [3] is that, as one of its tests of quality of knowledge, it attempts the kind of multiple choice examinations given to medical students as a major part of the process of satisfying medical licensing boards. These tests are undertaken by medical students to obtain a license to practice medicine, and as practice and self-assessment tests in preparation for taking these (see Section 1.9). Only secondarily is our project an investigation of how the same algorithms might be applicable to CDS, though this is potentially an important spin-off. As with real students, along with formally receiving knowledge, practicing these exams is part of the learning process. Similarly, MARPLE is told the official "correct answer", but only after it has made an attempt to answer the question, which contributes the official final exam score when we present results as if it were an exam taken by a medical student. Curation of the KRS to satisfy the criterion of good exam performance and learning from the examinations (as well as "learning" by receiving knowledge from data mining) become essentially the same thing. The level of automation of curation is already fairly high. There is only human intervention into the KRS when MARPLE persistently fails to answer an exam questions correctly. Before that happens, MARPLE takes considerable effort to seek out the knowledge required to answer the exam question, without human intervention.[2] By inspecting the question and most importantly the candidate answers, MARPLE queries the Internet and, having extracted knowledge from one web page, it explores more deeply by searching in turn on links found, including any found in a list of scientific references.

1.5. Further purposes of the present paper

Having a computer tackle medical school exams might seem likely to require fairly advanced techniques in Artificial Intelligence (AI). However, our preliminary studies [3] suggested that while such exams obviously test some important qualitative aspects of captured knowledge, they represent a rather restricted and "artificially crisp" world, to allow the student every chance to verify his or her knowledge. Our study is allowing us to comment of features that are of educational interest, as well a dissecting out some issues that do suggest some useful tools for CDS that are at least of the flavor of AI. For example, such exams still clearly test knowledge, but usually the reasoning with that knowledge almost always only requires *pre-syllogistic logic*, a term that we introduce in more detail in Section 5.1 but provide the theoretical basis in Sections 2.3 and 2.4. Apart from calculation questions that are excluded in the present study, very few exceptions to that have been found. Whereas more complete logic will be important for applications like CDS, it is a powerful pre-filter. It relates to the somewhat surprising finding that so-called *contextual probabilities* were sufficient [3]. These were originally intended just to be rough empirical estimates of *prior probabilities* for identifying the most likely answer, but they worked well with or without the support of the more precise kinds of *intrinsic probabilities* that are much more familiar, and which we would ideally like to test for CDS as well as flat statements of knowledge. These are probabilities as degrees of truth or scope intrinsically associated with each such statement. MARPLE 2 has a much improved ability to separate parts of the calculations and the parts of the KRS so that we can perform "computer experiments" giving insight into the above issues, as described in this paper. We also describe technological progress since MARPLE 1. It is by processes of Internet searching, preliminary

[2] The deep relationship between MARPLE learning from examinations (by ultimately noting the official correct answer) and simply acquiring knowledge from the Internet (irrespective of being tested in any examinations) is revealed by a thought experiment. We imagine the best case that a web page contains text that is essentially the question with the correct answer provided in the course of discussion. Indeed, published clinical case studies are essentially of that nature (Section 1.9). In practice, MARPLE does not often hit upon anything like that which is directly relevant. However, given long enough, and noting that hundreds of thousands of extracts of knowledge can easily be generated in this way in a day, the required information is likely to be found, even if element by element. Unfortunately, accumulating too much knowledge in the rougher XTRACT form contributes noise (Section 1.6).

Download English Version:

https://daneshyari.com/en/article/504958

Download Persian Version:

https://daneshyari.com/article/504958

Daneshyari.com