Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 298 (2016) 207–221

FUZZY sets and systems

www.elsevier.com/locate/fss

Reasoning with graded information: The case of diagnostic rating scales in healthcare

Thomas Vetterlein [a], Anna Zamansky [b,*]

[a] *Johannes Kepler University, Linz, Austria*
[b] *University of Haifa, Israel*

Received 12 October 2014; received in revised form 29 October 2015; accepted 3 November 2015

Available online 10 November 2015

Abstract

In medicine one frequently deals with vague information. As a tool for reasoning in this area, fuzzy logic suggests itself. In this paper we explore the applicability of the basic ideas of fuzzy set theory in the context of medical assessment questionnaires, which are commonly used, for instance, to support the diagnosis of psychological disorders. The items of a questionnaire are answered in a graded form; patients are asked to choose an element on a linear scale. The derived diagnostic hypotheses are graded as well. This leads to the question whether there is a logical formalism that is suitable to capture the score calculation of medical assessment questionnaires and thereby provides a mathematical justification of the way in which the calculation is typically done. We elaborate two alternative approaches to this problem. First, we follow the lines of mathematical fuzzy logic. For the proposed logic, which can deal with the formation of mean values, we present a Hilbert-style deduction system. In addition, we consider a variant of the prototype approach to vagueness. In this case we are led to a framework for which to obtain a logical calculus turns out to be difficult, yet our gain is a model that is conceptually comparably well-justifiable.

Keywords: Reasoning under vagueness; Healthcare applications; Assessment questionnaires; Approximate reasoning; Fuzzy logic

1. Introduction

Fuzzy set theory and fuzzy logic have originally been developed with the intention to overcome the particular difficulties that arise when an application requires the evaluation of *vague* information. Here, we call a property vague if it cannot be sharply delimited from its opposite. By default, a property that is applicable to a certain class of objects or processes is thought of as being dichotomous; it serves us to distinguish those cases in which it holds from those in which it does not hold. However, considering a situation in more detail, we may notice that it might not under all circumstances be possible to decide if a property applies or not. In fact, a characteristic feature of vague properties is the presence of borderline cases.

* Corresponding author.
E-mail addresses: thomas.vetterlein@jku.at (T. Vetterlein), annazam@is.haifa.ac.il (A. Zamansky).

http://dx.doi.org/10.1016/j.fss.2015.11.002

When wondering about examples of vagueness, there is one field that offers an apparently unlimited amount of notions that belong to this category. In fact, in medicine, it might, conversely, be found difficult to find descriptions of a different kind. A property describing the state of some patient is usually vague; we cannot use it positively or negatively under all possible circumstances. More specifically, we think about signs and symptoms of human diseases. Consider, e.g., the property of "having high fever". Below a temperature of, e.g., 38.5°, nobody would speak about high fever; above, e.g., 39.5° it is not questionable to speak about high fever; the remaining cases are borderline.

A common approach to deal with borderline cases is to extend the two-element set of truth values to a continuous-valued one; the two values "false" and "true" are replaced by the real unit interval. The property of having high fever, e.g., may be conveniently described by a fuzzy set mapping each possible temperature to such a generalised truth value. This is the very idea of fuzzy set theory [28], which, as far as the pure use of grades is concerned, is intuitively quite convincing.

We may ask to which extent the idea has been established as a tool for reasoning in medicine. The situation is somewhat ambiguous. Fuzzy logic has been applied, under this name, in the framework of several projects concerned with automated reasoning about medical information. In fact, in this context, its use seems to be clearly implied, given the fact that most of the processed information is vague. Already the system that is often called the "grandfather" of medical expert systems, MYCIN, was based on a continuous set of degrees [24]. Systems of this or a similar kind, however, are not often found in routine use in healthcare. To establish the principles of fuzzy logic in medicine – principles that are not unquestioned even in the community of fuzzy logicians – is certainly hard. At least in some cases, the basic ideas have been welcomed. For instance, the programming language Arden Syntax, designed for a platform-independent representation of medical knowledge, has been extended to simplify the treatment of fuzzy sets and is now called Fuzzy Arden Syntax [27]. Furthermore, MONI, a decision support system based on a simple fuzzy logic, is at present in use at the General Hospital of Vienna; its purpose is the automated detection of hospital-acquired infections [4].

Given the limited presence of fuzzy logic in medical decision support, the aspect that we address in the present paper provides a remarkable contrast. Grades are not just used within computer-based decision support systems; and clinicians do not necessarily consider grades as somewhat academic. In another context gradedness of information comes into play quite naturally.

The *assessment of symptoms* is an essential aspect of the diagnostic procedure for various disorders, in particular in psychiatry and psychology [26]. We recall that, as opposed to signs, which are objective phenomena detected by the clinician, symptoms are subjective experiences reported by the patient, such as a complaint of pain or depressed feeling. Thus, with symptoms, clinicians must rely on the patient's self-report, with no objective tests being available to confirm or rule out them [15]. This in turn opens up the possibility that patients report their symptoms autonomously. The use of questionnaires has contributed to a reduction of the working load in healthcare [14,25].

As a key element that we find in this context, the occurring questions can be considered as vague. In accordance with this observation, the answers are given in the form of grades. Consider, for instance, the ICD-10 Symptom Rating (ISR) questionnaire [25], which was created for the assessment of several psychological disorders. In order to evaluate the depressive syndrome, an answer to the item "I feel down and depressed" is required. To this end, the patient is asked to choose an element on a five-element linear scale, ranging from "0 – does not apply" to "4 – applies extremely". Apparently, we can understand the question as vague and the patient makes a choice to which *degree*, from his point of view, the indicated statement applies to his actual state.

The question on which the present work is based is now: how are the degrees further processed? Commonly, the answers to all questions are aggregated to a single value; questionnaires that are used in routine healthcare are commonly evaluated by the calculation of *rating scores*. Thus the question is which aggregation method is used for which reason. Ideally, questionnaires are designed and evaluated on the basis of well justified principles. In fact, within medical computer science, the topic has found an increasing interest during recent years. A large volume of works focuses on developing theories for a *statistical* analysis of questionnaire data, such as classical test theory or item response theory; see, e.g., [23,19,5].

These approaches might, on the one hand, provide useful insights into an optimal score calculation in questionnaires, considering, e.g., the weight and mutual dependency of items in a proper way. In practice, on the other hand, a procedure seems often to be chosen on pragmatic grounds. Consider, for instance, the development of the aforementioned ISR questionnaire. In this case, a panel of experts decided on the basis of their medical expertise. They voted on the method of calculating the total score and chose the mean value [25]. The experts also discussed the