Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 34 (2008) 2826–2840
Expert Systems with Applications
www.elsevier.com/locate/eswa

A Delphi-based approach to developing expert systems with the cooperation of multiple experts

Hui-Chun Chu, Gwo-Jen Hwang *

Department of Information and Learning Technology, National University of Tainan, 33, Sec. 2, Shulin Street, Tainan City 70005, Taiwan, ROC

Abstract

Knowledge acquisition has been a critical bottleneck in building knowledge-based systems. In past decades, several methods and systems have been proposed to cope with this problem. Most of these methods and systems were proposed to deal with the acquisition of domain knowledge from single expert. However, as multiple experts may have different experiences and knowledge on the same application domain, it is necessary to elicit and integrate knowledge from multiple experts in building an effective expert system. Moreover, the recent literature has depicted that "time" is an important parameter that might significantly affect the accuracy of inference results of an expert system; therefore, while discussing the elicitation of domain expertise from multiple experts, it becomes an challenging and important issue to take the "time" factor into consideration. To cope with these problems, in this study, we propose a Delphi-based approach to eliciting knowledge from multiple experts. An application on the diagnosis of Severe Acute Respiratory Syndrome has depicted the superiority of the novel approach.

Keywords: Knowledge acquisition; Expert system; Knowledge integration; Medical diagnosis; Delphi method

1. Introduction

In the past decades, expert systems have been applied to various applications. Subject domains that are supported by experts systems include bioengineering, defense, education, engineering, finance, and medical diagnosis. For example, MYCIN project is a well-known medical expert system for diagnosing infectious diseases (Buchanan & Shortliffe, 1985); ISODEPOR was developed to evaluate the muscle strength of Spanish top-competition athletes (Barreiro et al., 1997); FRBS-GP is a fuzzy rule-based system for diagnosing aphasia's subtypes and the classification of pap-smear examinations (Jantzen, Axer, & Keyserlingk, 2002).

The successful cases of the expert system approach not only demonstrated the benefits of applying expert system approach to coping with medical diagnosis problems, but also depicted the difficulty of applying it. In building an expert system, the critical bottleneck is to obtain the knowledge of the special domain from the domain experts, which is called *knowledge acquisition*. In past decades, several methods and systems have been proposed to cope with this problem. However, most of these methods and systems were proposed to deal with the acquisition of domain knowledge from single expert. However, as multiple experts may have different experiences and knowledge on the same application domain, it is necessary to elicit and integrate knowledge from multiple experts in building an effective expert system. Recent literature also indicated that "time" is an important parameter that might significantly affect the accuracy of inference results of an expert system; therefore, while discussing the elicitation of domain expertise from multiple experts, it becomes a much more challenging and important issue to take the "time" factor into consideration (Hwang, Chen, Hwang, & Chu, 2006).

To cope with these problems, we shall propose a Delphi-based approach to eliciting knowledge from multiple experts. An application of developing a medical expert system has depicted the superiority of the novel approach.

* Corresponding author. Tel.: +886 915396558; fax: +886 6 2606132.
E-mail address: gjhwang@mail.nutn.edu.tw (G.-J. Hwang).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.05.034

2. Relevant researches

To cope with the knowledge acquisition problem, many knowledge acquisition tools or methods have been proposed to build rapid prototypes and to improve the quality of the elicited knowledge, e.g., ETS (Boose, 1984, 1985), TEIRESIAS (Davis, 1979), MORE (Kahn, Nowlan, & McDermott, 1985), SALT (Marcus, 1987; Marcus, McDermott, & Wang, 1985), NeoETS (Boose & Bradshaw, 1986; Kitto & Boose, 1986), KNACK (Klinker, Bentolila, Genetet, Grimes, & McDermott, 1987), AQUINAS (Boose & Bradshaw, 1987; Shema & Boose, 1988), KRITON (Diederich, Ruhmann, & May, 1987), Student (Gale, 1987), RuleCons (O'Bannon, 1987), MOLE (Eshelman, Ehret, McDermott, & Tan, 1987), KITTEN (Shaw & Gaines, 1987), KSSO (Gaines, 1987), ASK (Gruber, 1988), WordNet (Millar, 1990; Navigli, Velardi, & Gangemi, 2003), KADS (Schreiber, Wielinga, & Breuker, 1993; Wielinga, Schreiber, & Breuker, 1992), MCRDR (Kang, 1996), MedFrame/CADIAG-IV (Boegl, 1997; Kolousek, 1997; Leitich et al., 2001). Most of these systems were developed based on the repertory grids method originated from Kelly's Personal construct theory (Kelly, 1955), which assists in identifying different objects in a domain and distinguishing among these objects.

A single repertory grid is represented as a matrix whose columns have elements labels and whose rows have construct labels. A 5-scale rating mechanism is usually used in filling the grid; i.e., each rating is an integer ranging from 1 to 5, where "1" represents that the element is very likely to have the trait; "2" represents the element may have the trait; "3" represents "unknown" or "no relevance"; "4" represents that the element may have the opposite characteristic of the trait; "5" represents that the element is very likely to have the opposite characteristic of the trait.

As repertory grid approach has been widely used by researchers, some extensions have been made to enhance its representative ability. For example, Jose, Nicholas, Jennings, Luo, and Shadbolt, 2003 developed a technique using a fuzzy repertory grid for acquiring the finite set of attributes or variables that the expert uses in a classification problem, characterizing and discriminating a set of elements. In addition, several models have been proposed to generate more meaningful rules from the repertory grid-oriented approaches, such as the EMCUD method, which can generate embedded meanings from repertory grids by defining the impacts of the constructs to each element (Hwang & Tseng, 1990). Recently, Hwang et al. (2006) indicated that, in building medical expert systems, most of the previous knowledge acquisition methods only pay attentions to the relationships between diseases and symptoms, while, the variant of the symptoms in different time scales of the diseases are not taken into account. Consider the repertory grid given in Table 1 which depicts an example of eliciting knowledge for diagnosing various kinds of gastrointestinal diseases. Note that the rating of the (Acute bronchitis, Throat pain) entry is 4, which implies highly tendency for Acute bronchitis to have Throat pain. However, in practical situation, Influenza has significant appearance of Throat pain in the early time scale. What has been addressed in the repertory grid is not happened in the last time scale of acute bronchitis. For later time scale, the throat pain symptom will become not so significant. Such variant of disease symptoms with respect to different time scales cannot be precisely presented by those conventional knowledge acquisition approaches.

3. Delphi-based knowledge acquisition approach

In developing a knowledge-based system, it is very difficult to elicit and integrate knowledge from multiple experts (Hwang et al., 2006), especially the application domains in which various time scales of elements need to be taken into account. To cope with this problem, a novel approach, Knowledge Acquisition for Multiple Experts with Time scales (KAMET), is proposed in this section, which takes time scales into consideration while eliciting expertise from multiple experts. In addition to time scales, KAMET takes importance degree for each construct to each element in different time scales into account, such that more embedded knowledge can be explicitly presented.

Let e_i^t denote tth stage period of element (or disease) e_i and c_j denote a construct (or symptom), where $i = 1$ to n, and $j = 1$ to m. Each KAMET entry is a triplet that consists of three values: a rating to indicate the relevance of e_i^t and c_j, a certainty degree for giving the rating and an impact factor to represent the importance of c_j to e_i^t, which are represented by the following three functions:

(1) Rating (e_i^t, c_j): the degree of relevance for element e_i in tth time scale to construct c_j, ranging from 1 to 5: "1" represents that the element is very likely to have the opposite characteristic of the trait; "2" represents the element may have the opposite characteristic of the trait; "3" represents "unknown" or "no relevance"; "4" represents that the element may have the trait; "5" represents that the element is very likely to have the trait.
(2) Certainty (e_i^t, c_j): the degree of certainty for giving Rating (e_i^t, c_j), which is either "S" or "N" representing "sure" or "not sure".
(3) Impact_factor (e_i^t, c_j): the degree of importance for construct c_j to element e_i in tth time scale. Impact_factor (e_i^t, c_j) can be one of the following values: "*X*" represents no relationship between the element and the construct; "*D*" means that the construct dominates the element, i.e., if the value of the construct is not matched, it is impossible for the element to be implied; an integer, ranging from 1 to 5, indicates that the construct is of some degree of importance to the element, but does not dominate the implication of the element.