Knowledge-Based Systems 59 (2014) 121–131

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Elicitation criterions for restricted intersection of two incomplete soft sets

Bang-He Han [a,*], Yong-Ming Li [b], Jie Liu [c], Sheng-Ling Geng [d], Hou-Yi Li [e]

[a] School of Mathematics and Statistics, Xidian University, Xi'an 710071, China
[b] College of Computer Science, Shaanxi Normal University, Xi'an 710062, China
[c] School of Computer Science and Technology, Xidian University, Xi'an 710061, China
[d] College of Computer, Qinghai Normal University, Xining 810008, China
[e] State Key Laboratory of ISN, School of Telecommunications Engineering, Xidian University, Xi'an 710061, China

ARTICLE INFO

Article history:
Received 18 March 2013
Received in revised form 14 January 2014
Accepted 18 January 2014
Available online 31 January 2014

Keywords:
Elicitation
Incomplete soft set
Restricted intersection
Decision making
Choice value

ABSTRACT

Purpose: This paper aims to develop and compare several elicitation criterions for decision making of incomplete soft sets which are generated by restricted intersection.
Design/methodology/approach: One time elicitation process is divided into two steps. Using the greedy idea four criterions for elicitation of objects are built based on maximax, maximin, minimax regret and combination of expected choice values and elicitation times. Then these initial unknown values which produce incomplete values together with known information are in priority.
Findings: Fast methods for computing possibly and necessarily optimal solutions before or in the elicitation process are invented. As far as the sizes of soft sets used in the simulation experiments, it is found statistically that we should choose the criterion based on the combination of expected choice value and expected elicitation times in the first step of one time elicitation.
Practical implications: The developed methods can be used for decision making of incomplete 0–1 information systems, which are generated by the conjunction of two experts' incomplete 0–1 evaluation results. Whenever the available information is not enough for choosing a necessarily optimal solution, the elicitation algorithms can help elicitate as few unknown values as possible until an optimal result is found. An elicitation system is made to show that our elicitation methods can potentially be embedded in recommender or decision support systems.
Originality/value: The elicitation problems are proposed for decision making of operation-generated soft sets by extracting from some practical problems. The concept of expected elicitation times of objects is defined and used for developing one type of elicitation strategy.

1. Motivations

This section mainly focuses on the motivations of this paper's subject. Firstly two cases in practice are given. Then we make an introduction to soft sets. At last by describing the problems in the two cases with soft set model, we propose the elicitation problems to be dealt with in this paper.

1.1. Two cases in real practice

Case 1: Site Selection

A company leader wants to make a site selection for his overseas branch company. $U = \{s_1, s_2, s_3, s_4\}$ is the set of 4 potential choices. $E = \{c_1, c_2, c_3, c_4\}$ is the set of parameters describing conditions such as air humidity, traffic facilities, price of land, labor price. There are two investigators Alice and Bob. They have made their incomplete evaluations which are shown in Table 1, where.

- in each bracket:

(1) the first value refers to the opinion of Alice and the second value refers to that of Bob.

* Corresponding author. Tel.: +86 29 81891371.
E-mail addresses: bhhan@mail.xidian.edu.cn (B.-H. Han), liyongm@snnu.edu.cn (Y.-M. Li), 835088577@qq.com (J. Liu), qhnugsl@gmail.com (S.-L. Geng), lihouyi2008@126.com (H.-Y. Li).

http://dx.doi.org/10.1016/j.knosys.2014.01.015
0950-7051/

Table 1
Integration of evaluation matrixes from Alice and Bob by min operator.

	c_1	c_2	c_3	c_4	σ
s_1	1(1,1)	0(0,0)	0.5(0.5,1)	1(1,1)	$2+1^*$
s_2	0(0,1)	0(1,0)	1(1,1)	1(1,1)	2
s_3	1(1,1)	0.5(0.5,0.5)	1(1,1)	0(0,0)	$2+1^{**}$
s_4	1(1,1)	0.5(0.5,1)	0.5(0.5,0.5)	0(0,1)	$1+1^*+1^{**}$

(2) 1 means appropriate (positive) and 0 means a negative answer.
(3) 0.5 means incomplete evaluation.

- the value outside each bracket is equal to the minimum of the two values inside the bracket. This means that a certain choice is appropriate with respect to a certain condition if and only if the two investigators are both satisfied with that. In the situation when Alice is satisfied with certain condition of one choice and Bob has not given his opinion on this, the integration of their evaluations should be incomplete. Obviously, once Alice says no for a certain condition of one choice, the final integration must be negative no matter what the opinion of Bob is.
- the value of form $a+b^*+c^{**}$ in the row s_i ($i = 1,2,3,4$) and the column σ corresponds to the incomplete decision value of s_i :

(4) a equals to the number of value 1 outside brackets in the row s_i.
(5) Definition 1.1. b^* is defined as the incomplete decision value part marked with one asterisk of s_i, where b equals to the number of value 0.5 outside these brackets containing only **one** 0.5 in the row s_i. When $b=0, b^*$ is omitted.
(6) Definition 1.2 c^{**} is defined as the incomplete decision value part marked with two asterisks of s_i, where c equals to the number of value 0.5 outside these brackets containing **two** 0.5 in the row s_i. When $c=0, c^{**}$ is omitted.

For Table 1, the decision value of s_2 is equal to 2. However, the real decision values of the other objects are unavailable due to the incomplete information. In this example all of these 4 choices are possible to one of the optimal solutions under usual additive model and no one of them must be an optimal solution of the real situation. If we want to make approximate reasoning, which one should be an approximate solution? If we want to make exact reasoning, which investigator should give evaluation on which unknown value?

Case 2: Finding a best employee

A company wants to employ one person from a lot of graduates. There are two interviewers to evaluate these job hunters about the following aspects: c_1, writing ability; c_2, organizing ability; c_3, communicating ability; c_4, adaptability to changes. The evaluation results of the two interviewers can be given by two 0–1 valued tables, where the rows correspond to the job hunters, the columns correspond to the above four abilities. 1 means good, 0 means not good. The final evaluation matrix is got by min operator similarly like **Case 1**. For certain reasons, these evaluation tables may be incomplete. Table 1 can also be an example of this case, where s_1 to s_4 mean four graduates. In this situation the job interviewers have to make efforts to know some of these unknown values. So who should do elicitation work? And about which unknown value? We need an elicitation algorithm for guiding this work.

1.2. An introduction to soft set

Soft set theory was initiated as a new mathematical tool for dealing with uncertainty and vagueness by Molodtsov [27]. The theory of soft sets [24,26] has potential wide applications in fields like game theory, operations research, decision making and so on. Many researchers have studied soft set based decision making [4,11,14,15,19–23,25,30,32,34]. For instances, soft sets theory have been used to evaluate business competitive capacity [37] and make combined forecasting approach [38].

Table 2
Tabular representation of a (an incomplete) soft set P with (incomplete) decision values.

	e_1	e_2	e_3	e_4	e_5	σ
h_1	0	1	0	1	0	2
h_2	1	0	1(0.5)	0	1	3(2 + 1*)
h_3	1	1(0.5)	0	0	1	3(2 + 1*)
h_4	0	1	1	1	1	4

Let us give an example of a soft set. Assume that $U=\{h_1,h_2,h_3,h_4\}$ is a set of houses. $A=\{e_1,e_2,e_3,e_4,e_5\}$ is a set of parameters characterizing "cheap", "legal", "subway will be build nearby", "convenient for shopping", "near school", respectively. A soft set P over U is a function from A to the powerset of U. That is for each $e_i \in A$, there exists a subset of U, which means that the houses in this subset is good with respect to this parameter. See Table 2 for a representation of a soft set P over U. The subset corresponding with e_1 is equal to $\{h_2,h_3\}$. The subset corresponding with e_2 is equal to $\{h_1,h_3,h_4\}$.

The decision value function of this soft set, denoted by σ, compute for each object h_i the number of parameters whose corresponding subset of U contains h_i. From the point view of Table 2, $\sigma(h_i)$ is the number of value 1 in the row indexed by h_i. According to soft set theory, the best house is the one which has the maximum value of σ.

However, soft set may be incomplete. Take the soft set in Table 2 for example, the businessman in real estate may cheat customer or refuse to show the required certificates. Thus people have to make some efforts to consult relative government departments. Similarly whether there will be subway built nearby, for some houses it is hard to say. We have to consult the municipal construction planning.

1.3. Elicitation problems in soft set theory

By comparing Tables 1 and 2, it is easy to see that the tabular representations of evaluation matrixes of Alice and Bob in **Case 1** can be regarded as incomplete soft sets. A complete evaluation from Alice or Bob can both be regarded as complete soft sets. Similarly for **Case 2**. What is more, the integration method can actually be modeled by the restricted intersection operation of two soft sets.

As an abstraction of the questions in **Case 1** nd **Case 2** in the context of soft set theory, we assume such a situation in which the unknown information of incomplete soft sets can be accessed under certain conditions. For example, we can pay for that. We are to find at least one solution which must be an optimal solution of the real soft set no matter what the missing values are. So we need an elicitation algorithm. Simply speaking, we need to answer which unknown value to elicitate first. Such algorithm should be better in the sense that we need to elicitate as few unknown values as possible. Elicitation strategies for soft constraint problems with missing preferences have been studied by Gelain et al. [17] and Pini et al. [29]. The difference is that we are dealing with elicitation problems of incomplete 0–1 information systems [28] which are generated by another two ones.

Why we choose soft set model? The reasons are as follows.

- In real practice, Alice and Bob may evaluate different set of parameters. And the integration method may have some other models rather than the min operator (conjunction). In soft set theory these situations have already been