Information Sciences 346–347 (2016) 424–441

Contents lists available at ScienceDirect

Information Sciences

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

Selected group-theoretic aspects of confirmation measure symmetries

Robert Susmaga*, Izabela Szczęch

Institute of Computing Science, Poznań University of Technology, Piotrowo 2, 60–965 Poznań, Poland

ARTICLE INFO

Article history:
Received 10 July 2014
Revised 19 October 2015
Accepted 20 January 2016
Available online 4 February 2016

Keywords:
Bayesian confirmation measures
Symmetry properties
Elements of the group theory

ABSTRACT

The paper considers symmetry properties of rule interestingness measures (in particular: measures of confirmation). Many authors have studied various symmetries, however the discussion on which sets of symmetries should be taken into account, let alone why the particular symmetries are desirable or not, has still not produced a generally recognized consensus. Furthermore, the results published so far neglect the fact that symmetries can be the subject of group theory-based considerations. This paper aims at solving those problems by introducing group-theoretic interpretations of symmetries, indicating that all symmetries can be treated as permutations, and compositions of symmetries as compositions of permutations. Such an interpretation allows us to apply the well-known group-theoretic results to symmetries. In particular, using this approach, we reveal the phenomenon of incompleteness occurring in sets of symmetries considered by different authors, and propose an effective way of controlling it. Moreover, we demonstrate that assessing the symmetries as either desirable or undesirable, as introduced by these authors, brings in inconsistencies, which become evident under the permutation-based interpretation. Finally, we present group-theoretic guidelines to the design of such symmetry sets that are free of the incompleteness and inconsistency phenomena but remain meaningful in the context of rule evaluation.

1. Introduction

The evaluation of "*if* premise, *then* conclusion" rules (denoted as $E \rightarrow H$) induced from data is commonly done by means of interestingness measures. The measures are used to filter out rules which are of low utility for the user. In [12], Greco et al. have shown that also measures known as Bayesian confirmation measures (sometimes referred to as measures of evidential support [9]) can act as interestingness measures for the decision rules. Indeed, Bayesian confirmation measures have been used in many practical applications e.g., in rule-discovery systems [10,21,23] or recently in bio-medical systems [18]. In particular, they have played a major, albeit rarely acknowledged, role in the renowned expert system MYCIN (the so called certainty factor, which is a fundamental coefficient used by this system, happens to be equivalent to Bayesian confirmation measure $Z(H, E)$ [5]).

Despite the fact that the class of Bayesian confirmation measures is quite large, all of its members have one common feature – they quantify "to what extent the rule's premise confirms or disconfirms the rule's hypothesis". Let us observe that in the literature there are many ordinally non-equivalent measures of Bayesian confirmation (which means that they

* Corresponding author. Tel.: +48 618790790; fax: +48 618771525.
E-mail addresses: robert.susmaga@cs.put.poznan.pl (R. Susmaga), izabela.szczech@cs.put.poznan.pl (I. Szczęch).

http://dx.doi.org/10.1016/j.ins.2016.01.041
0020-0255/

produce different rule rankings). In order to discriminate between them and eventually help the user choose one for a particular application, numerous properties of measures have been discussed and studied. In particular, this paper focuses on symmetry properties of these measures.

There are many studies concerning these properties (see [4–7,13,14]), differing in the set of considered symmetries or in the context in which the analysis is conducted (context-free or context-based, the latter incorporating the information about situation of confirmation or disconfirmation). Moreover, properties of symmetry have been differently assessed as either desirable (advantageous) or undesirable (disadvantageous) by many authors, and there is still no general consensus on such a preferential assessment. This is of great loss for the user who wishes to choose a measure responsibly, but is not too well guided to any particular set of desirable symmetries.

In the context of these problems, we propose to look at symmetry properties from a new perspective and to consider them as the subject of the group theory. In particular, we provide group-theoretic interpretations of the symmetries, indicating that all symmetries can be treated as permutations, and compositions of symmetries as compositions of permutations. Application of such a group-theoretic perspective allows to identify incompleteness phenomena in particular sets of symmetries considered by different authors. However, as we demonstrate, enriching the sets of symmetries discussed in the literature with some group-theoretic properties allows to eliminate the cases of incompleteness. In particular, the paper shows that supplementing the set of symmetries considered in [5,14] with a trivial element changes it into a group that is isomorphic to the dihedral group D_4 (rigid transformations of a square), and therefore free of the incompletenesses.

The group-theoretic results are also very beneficial with respect to assessing which symmetries should be considered as desirable and which as undesirable. In particular, the group-theoretic interpretation of symmetries reveals inconsistencies in preferential assessments proposed by some authors. As a way of controlling such inconsistencies, the paper presents postulates that may be used to justify the partitioning of the symmetries into consistent sets of desirable and undesirable ones.

The obtained results are thoroughly discussed, helping reach an agreement as to which symmetries are ultimately desirable for confirmation measures.

The rest of the paper is organized as follows. Section 2 describes examples of popular confirmation measures, lays out the concept of Bayesian confirmation, and presents the different sets of measure symmetries together with their preferential assessment, as introduced by different authors; it also introduces the permutation interpretation of these symmetries. Section 3 presents the required elements of the group theory, in particular the dihedral group of rigid transformations of a square, together with its most important properties. Section 4 introduces a group derived from the Bayesian confirmation conditions, demonstrating a clear correspondence between the dihedral group and the symmetries of confirmation measures. Section 5 introduces the concept of incompleteness in sets of symmetries and discusses the group-theoretic approach to incompleteness control. Section 6 introduces the concept of inconsistency in preferential assessments of symmetries and discusses two group-theoretic approaches to inconsistency control. Section 7 provides additional interpretation to selected results from the two previous sections, placing them in the context of rule evaluation. Final remarks and conclusions are contained in Section 8.

2. Bayesian confirmation measures

The paper concentrates on a group of interestingness measures called Bayesian confirmation measures (for a review on Bayesian confirmation measures see [3,5,8,12]). According to Fitelson [9], such measures quantify the degree to which the evidence in the rule's premise E provides support *for* or *against* the hypothesized piece of evidence in the rule's conclusion H.

Originally, the definitions of Bayesian confirmation measures were expressed using probabilities, see e.g. measures $D(H, E)$ [6] and $C(H, E)$ [4]:

$$D(H, E) = P(H|E) - P(H),$$

$$C(H, E) = P(E \wedge H) - P(E)P(H).$$

However, as far as the real-world phenomena described by the rules are concerned, specific information on whether a given piece of evidence E or hypothesis H holds or not is often estimated with four discrete, non-negative values:

- a: the number of objects in the dataset for which both E and H hold,
- b: the number of objects in the dataset for which the premise E does not hold, but the conclusion H holds,
- c: the number of objects in the dataset for which the premise E holds, but the conclusion H does not hold,
- d: the number of objects in the dataset for which neither E nor H holds.

These values, used throughout the paper, may collectively be stored in a 2 × 2 table, referred to as the contingency table (see Table 1), with rows and columns characterizing the premise and the conclusion, respectively. Alternatively, they will be briefly presented as a 2 × 2 matrix $[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}]$. The sum of these values will be in all cases denoted by n (so $n = a + b + c + d$).

Let us observe that Table 1 can be used to estimate probabilities of E, H, $\neg E$ and $\neg H$: e.g. the probability of the premise E is expressed as $P(E) = (a + c)/n$, the probability of the conclusion H as $P(H) = (a + b)/n$, and so on. Using such a