Analysis of symmetry properties for Bayesian confirmation measures

Salvatore Greco

University of Catania, Italy

Roman Słowiński

Poznań University of Technology,

Systems Research Institute*, PAS, Poland*

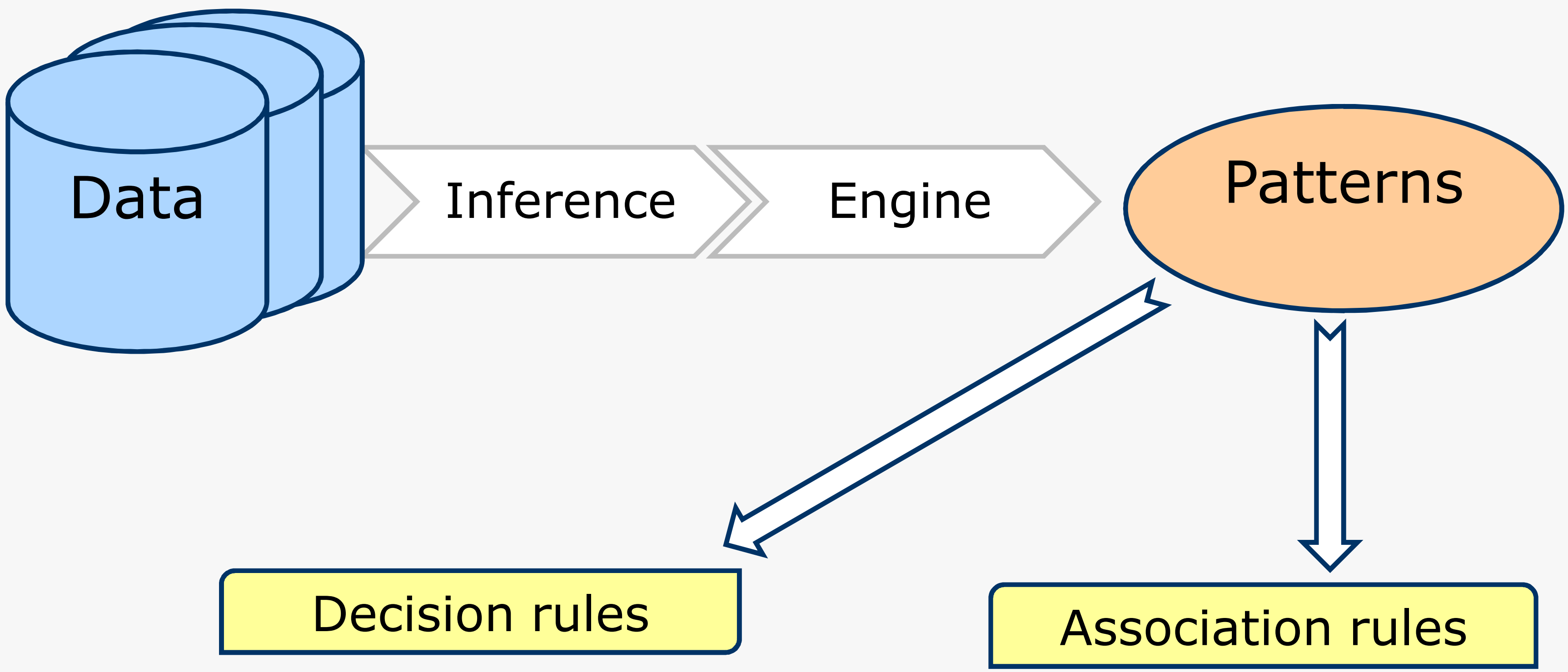
Izabela Szczęch

Poznań University of Technology, Poland

Presentation plan

- Rule induction and evaluation
- Bayesian confirmation measures
- Symmetry properties in literature
- Proposition of a new set of symmetry properties
- Conclusions

Rule induction



If symptom s_1 is present
and symptoms s_2
and s_3 are absent
then disease d_1

If symptom s_1 is present
then symptoms s_2
and s_3 are absent

Rule induction

- Patterns in form of rules are induced from a data table
- $S=\langle U, A\rangle$ – *data table*, where U and A are finite, non-empty sets
  U – *universe of objects*; A – set of *attributes*
- $S=\langle U, C, D\rangle$ – *decision table*, where C – set of *condition attributes*,
  D – set of *decision attributes*, $C\cap D=\emptyset$

- *Rule* induced from S is a *consequence relation*:
  E* → *H *read as* ***if E then H***
  where E is condition (evidence or premise)
  and H is conclusion (hypothesis or decision)
  formula built from attribute-value pairs (q,v)

Rule induction

Characterization of nationalities

U	*Height*	*Hair*	*Eyes*	*Nationality*	Support
1	*tall*	*blond*	*blue*	*Swede*	270
2	*medium*	*dark*	*hazel*	*German*	90
3	*medium*	*blond*	*blue*	*Swede*	90
4	*tall*	*blond*	*blue*	*German*	360
5	*short*	*red*	*blue*	*German*	45
6	*medium*	*dark*	*hazel*	*Swede*	45
		C		*D*	

- E.g. decision rules induced from „characterization of nationalities”:
  1) If (*Height=tall*), then (*Nationality=Swede*)
  2) If (*Height=medium*) & (*Hair=dark*), then (*Nationality=German*)

Interestingness measures

The number of rules
induced from datasets is usually quite large

- overwhelming for human comprehension,
- many rules are irrelevant or obvious
  (low practical value)

rule evaluation – interestingness (attractiveness) measures
(e.g. support, confidence, gain, rule interest, lift
measures of Bayesian confirmation)

- each measure was proposed to capture
different characteristics of rules
- the number of proposed measures is very large

Property of Bayesian confirmation

- An attractiveness measure $c(H,E)$, has the property of Bayesian confirmation (i.e. is a confirmation measure) if is satisfies the following condition:

$$c(H,E)\begin{cases} >0 \ \ if \ \ Pr(H|E) > Pr(H) \\ =0 \ \ if \ \ Pr(H|E) = Pr(H) \\ <0 \ \ if \ \ Pr(H|E) < Pr(H) \end{cases} \qquad c(H,E)\begin{cases} >0 \ \ if \ \ Pr(H|E) > Pr(H|\neg E) \\ =0 \ \ if \ \ Pr(H|E) = Pr(H|\neg E) \\ <0 \ \ if \ \ Pr(H|E) < Pr(H|\neg E) \end{cases}$$

- Measures of confirmation quantify the strength of confirmation that premise *E* gives to conclusion *H*
- „*H* is verified more often, when *E* is verified, rather than when *E* is not verified”

Bayesian confirmation measures

- The condition does not put any constraint on the value to be assigned to confirmatory arguments (as long as they are positive) or disconfirmatory arguments (as long as they are negative)

$$c(H,E)\begin{cases} >0 \ \ if \ \ Pr(H|E) > Pr(H|\neg E) \\ =0 \ \ if \ \ Pr(H|E) = Pr(H|\neg E) \\ <0 \ \ if \ \ Pr(H|E) < Pr(H|\neg E) \end{cases}$$

- There are many alternative, non-equivalent measures of Bayesian confirmation

- Used notation corresponding to a 2x2 contingency table of rule's premise and conclusion

	H	$\neg H$	Σ
E	a	c	$a+c$
$\neg E$	b	d	$b+d$
Σ	$a+b$	$c+d$	$a+b+c+d=\|U\|$

Popular measures of Bayesian confirmation

There are many alternative, non-equivalent measures of Bayesian confirmation

$$D(H,E)=\frac{a}{a+c}-\frac{a+b}{|U|}$$ (Carnap 1950/1962)

$$S(H,E)=\frac{a}{a+c}-\frac{b}{b+d}$$ (Christensen 1999)

$$M(H,E)=\frac{a}{a+b}-\frac{(a+c)}{|U|}$$ (Mortimer 1988)

$$N(H,E)=\frac{a}{a+b}-\frac{c}{c+d}$$ (Nozick 1981)

$$C(H,E)=\frac{a}{|U|}-\frac{(a+c)(a+b)}{|U|^2}$$ (Carnap 1950/1962)

$$R(H,E)=\frac{a|U|}{(a+c)(a+b)}-1$$ (Finch 1960)

$$G(H,E)=1-\frac{c|U|}{(a+c)(c+d)}$$ (Rips 2001)

$$F(H,E)=\frac{ad-bc}{ad+bc+2ac}$$ (Kemeny and Oppenheim 1952)

Utility of confirmation measures *F* vs. utility of confidence

- Consider the possible result of rolling a die: 1,2,3,4,5,6, and let the premise be kept fixed at *E="the result is divisible by 2"*
  - H_1 ="the result is 6" $conf(H_1, E)=1/3$, $F(H_1, E)=3/7$
  - H_2 ="the result is *not* 6" $conf(H_2, E)=2/3$, $F(H_2, E)=-3/7$

- In this example, rule $E \rightarrow H_2$ has greater confidence than rule $E \rightarrow H_1$
- However, rule $E \rightarrow H_2$ is less interesting than rule $E \rightarrow H_1$ because *premise E reduces the probability of conclusion* H_2 from $5/6=sup(H_2)$ to $2/3= conf(H_2, E)$, while it augments the probability of conclusion H_1 from $1/6=sup(H_1)$ to $1/3= conf(H_1, E)$
- In consequence, *premise E disconfirms conclusion* H_2, which is expressed by a negative value of $F(H_2, E)=-3/7$, and it confirms conclusion H_1, which is expressed by a positive value of $F(H_1, E)=3/7$

Properties of Bayesian confirmation measures

The choice of an interestingness measure for a certain application is a difficult problem

- there is no evidence which measure(s) is the best
- the users' expectations vary,
- the number of proposed measures is overwhelming

properties of interestingness measures, which reflect users' expectations towards the behavior of measures in particular situations

- property of monotonicity M (Greco, Pawlak & Slowinski 2004)
- Ex_1 property and its generalization to weak Ex_1
- property of logicality L and its generalization to weak L (Greco, Slowinski & Szczech 2012)
- …

need to analyze measures with respect to their properties

In this work we focus on a group of **symmetry properties**

Symmetry properties

- Symmetry properties are formed by applying the negation operator to the rule's premise/conclusion, or both, as well as switching the position of the premise and the conclusion.

- Carnap, Eells and Fitelson have analyzed confirmation measures from the viewpoint of four properties of symmetry
  - evidence symmetry *ES*: $c(H, E) = -c(H, \neg E)$
  - hypothesis symmetry *HS*: $\boldsymbol{c(H, E) = -c(\neg H, E)}$
  - commutativity (inversion) symmetry *IS*: $c(H, E) = c(E, H)$
  - total (evidence-hypothesis)symmetry *EHS*: $c(H, E) = c(\neg H, \neg E)$
- Their conclusion: only hypothesis symmetry *HS* is a desirable property

Evidence symmetry

- According to Eells and Fitelson
  evidence symmetry *ES*: $c(H, E) = -c(H, \neg E)$ should be discarded
- Let us consider a rule:

if the drawn card is the seven of spades then the card is black

  - the *seven of spades* confirms that the *card is black* to a greater extent than the *not-seven of spades* disconfirms the same hypothesis
  - thus, the equality in *ES* is found unattractive by Eells and Fitelson, i.e. for some situation $c(H, E) \neq -c(H, \neg E)$

Crupi et al. symmetries

- Recently, Crupi, Tentori and Gonzalez propose to analyze a confirmation measure $c(H, E)$ with respect to the following symmetries
  - $ES(H, E)$: $c(H, E) = -c(H, \neg E)$
  - $HS(H, E)$: $c(H, E) = -c(\neg H, E)$
  - $EIS(H, E)$: $c(H, E) = -c(\neg E, H)$
  - $HIS(H, E)$: $c(H, E) = -c(E, \neg H)$
  - $IS(H, E)$: $c(H, E) = c(E, H)$
  - $EHS(H, E)$: $c(H, E) = c(\neg H, \neg E)$
  - $EHIS(H, E)$: $c(H, E) = c(\neg E, \neg H)$
- Crupi et al. claim that the analysis should be conducted separately for:
  - the case of confirmation (i.e. when $\Pr(H|E) > \Pr(H)$), and
  - for the case of disconfirmation (i.e. when $\Pr(H|E) < \Pr(H)$)
- Such approach results in 14 symmetry properties

Crupi et al. symmetries – inversion symmetry

- Crupi et al. concur with the results of Eells and Fitelson regarding the inversion symmetry only in case of confirmation
- Crupi et al. claim that *IS* is desirable in case of disconfirmation
- Let us consider a rule:

  if the drawn card is an Ace, then it is a face

  - the strength with which an *Ace* disconfirms *face* is the same as the strength with which the *face* disconfirms an *Ace*, i.e. $c(H, E) = c(E, H)$
- Conclusions of Crupi et al.:
  - in case of confirmation only the *HS*, *HIS* and *EHIS* are the desirable properties
  - in case of disconfirmation only *HS*, *EIS* and *IS* properties are the desirable properties

A new set of symmetry properties

- Let us observe that the approaches of Eells and Fitelson as well as Crupi et al. mainly concentrate on entailment and refutation of the hypothesis by the premise
- This, however, boils the concept of confirmation down only to situations:
  - where there are no counterexamples (entailment) and
  - where there are no positive examples to a rule (refutation)
- A confirmation measure should give an account of the credibility that *it is more probable to have the conclusion when the premise is present, rather than when the premise is absent*
- Both conditional probabilities $\Pr(H|E)$ and $\Pr(H|\neg E)$ should be considered both in case of confirmation and disconfirmation
- There is no need to treat case of confirmation and disconfirmation separately

A new set of symmetry properties – evidence symmetry

- *ES*: $c(H, E) = -c(H, \neg E)$ is desirable
- Let us examine both sides of this equation using an exemplary scenario where the values of contingency table of *E* and *H* are:

	H	$\neg H$
E	$a=100$	$c=0$
$\neg E$	$b=99$	$d=1$

- Let us observe, that for $c(H, E)$ we have that
  $\Pr(H|\neg E)= b/(b+d)=0.99$ and
  $\Pr(H|E)=a/(a+c)=1$, which gives us a 1% increase of confirmation
- On the other hand, for $c(H, \neg E)$ we get:
  $\Pr(H|E)=1$ and $\Pr(H|\neg E)=0.99$, which results in 1% decrease

- Thus, clearly the confirmation of a rule $E \rightarrow H$ should be of the same value but of the opposite sign as the confirmation of a $\neg E \rightarrow H$ rule

A new set of symmetry properties – summary

ES	YES for any (H,E) $c(H, E) = -c(H, \neg E)$
HS	YES for any (H,E) $c(H, E) = -c(\neg H, E)$
EIS	NO for some (H,E) $c(H, E) \neq -c(\neg E, H)$
HIS	NO for some (H,E) $c(H, E) \neq -c(E, \neg H)$
IS	NO for some (H,E) $c(H, E) \neq c(E, H)$
EHS	YES for any (H,E) $c(H, E) = -c(\neg H, \neg E)$
EHIS	NO for some (H,E) $c(H, E) \neq -c(\neg E, \neg H)$

Conclusions

- Bayesian confirmation measures constitute a group of important and useful interestingness measures
- To help to handle the plurality of measures, their properties are analyzed
- We focused on a group of symmetry properties
- Our analysis was conducted regarding that confirmation measures should reflect *how much more it is probable to have the conclusion when the premise is present rather than when it is absent*
- Such interpretation of the confirmation concept leads to proposition of a new set of desirable symmetry properties: *ES*, *HS* and *EHS*
- Thus, valuable confirmation measures should only satisfy *ES*, *HS* and *EHS*

Future research

- Consequently, our future research will concentrate on verification which of the commonly used confirmation measures satisfy *ES*, *HS* and *EHS*, not enjoying the other symmetries at the same time
- Moreover, experiments on real datasets shall be performed to show the advantages of using such measures

Thank you!

Utility of Bayesian confirmation measures

- Using the quantitative confirmation theory for data analysis allows to benefit from the ideas of such prominent researchers as *Carnap*, *Hempel* and *Popper*
- Prof. *Pawlak* advocated that the group of Bayesian confirmation measures should be considered a valuable and meaningful tool for assessing the quality of rules induced from data, eg. within rough set approach and, more generally, within knowledge discovery
- By using the interestingness measures that possess this property one can filter out rules which are misleading and disconfirm the user