Tato prezentace je součástí wiki-prezentace Metoda GUHA a systém LISp-Miner
Je dostupná z této adresy
Verse 3. 9. 2019

Slabě implikační kvantifikátory a pravidla

Jan Rauch
Katedra informačního a znalostního inženýrství
Vysoká škola ekonomická v Praze

© Jan Rauch

Slabě implikační kvantifikátory a pravidla

- 4ft-kvantifikátor $\rightarrow_{C,S}$
- 4ft-kvantifikátor $\rightarrow_{C,S}$ není implikační - příklad
- Slabě implikační kvantifikátory a pravidla -úvaha
- Slabě implikační kvantifikátory a pravidla - definice
- Každý implikační kvantifikátor je slabě implikační
- Důležité slabě implikační kvantifikátory v proceduře 4ft-Miner

4ft-kvantifikátor $\rightarrow_{C,S}$

4ft-kvantifikátor $\rightarrow_{C,S}$ je definovaný tak, že asociační pravidlo $\varphi \approx \psi$ je pravdivé v matici dat **M** pokud pro čtyřpolní tabulku 4ft(φ, ψ, **M**) = $\langle a, b, c, d \rangle$ platí

- $\rightarrow_{C,S}(a, b, c, d) = 1$ pokud $\frac{a}{a+b} \geq C \wedge \frac{a}{a+b+c+d} \geq S$
- $\rightarrow_{C,S}(a, b, c, d) = 0$ jindy.

C je konfidence a S je podpora asociačního pravidla $\varphi \approx \psi$, viz též úvod

Čtyřpolní tabulku 4ft(φ, ψ, **M**) zapisujeme také takto:

M	ψ	$\neg\psi$
φ	a	b
$\neg\varphi$	c	d

4ft-kvantifikátor $\rightarrow_{C,S}$ není implikační - příklad

Aby byl 4ft-kvantifikátor $\rightarrow_{C,S}$ implikační, musí pro čtveřice $\langle a, b, c, d \rangle$ a $\langle a', b', c', d' \rangle$ celých nezáporných čísel platit:

jestliže $\rightarrow_{C,S} (a, b, c, d) = 1$ a zároveň $a' \geq a$ a $b' \leq b$, pak i $\rightarrow_{C,S} (a', b', c', d') = 1$.

Použijeme 4ft-kvantifikátor $\rightarrow_{0.9,0.1}$, a čtveřice

$\langle a, b, c, d \rangle = \langle 90, 10, 0, 0 \rangle$ a $\langle a', b', c', d' \rangle = \langle 100, 0, 9900, 0 \rangle$. Platí:

- $\rightarrow_{C,S} (90, 10, 0, 0) = 1$ protože $\frac{90}{90+10} \geq 0.9 \wedge \frac{10}{90+10+0+0} \geq 0.1$
- $a' \geq a$ a $b' \leq b$ protože $100 \geq 90$ a $0 \leq 10$
- $\rightarrow_{C,S} (100, 0, 9900, 0) = 1$, protože sice $\frac{100}{100+0} \geq 0.9$, ale

$$\frac{100}{100 + 0 + 9900 + 0} = \frac{100}{10000} = 0.01 < 0.1$$

Slabě implikační kvantifikátory a pravidla - úvaha

- Zajímají nás korektní dedukční pravidla $\frac{\varphi\approx\psi}{\varphi'\approx\psi'}$. Chceme, aby z platnosti $\varphi\approx\psi$ na **M** plynula platnost $\varphi'\approx\psi'$ na **M** pro každou matici dat **M**.
- Tedy, aby z platnosti $\approx(a, b, c, d) = 1$ plynulo $\approx(a', b', c', d') = 1$, kde $\langle a, b, c, d \rangle = 4\mathrm{ft}(\varphi, \psi, \mathbf{M})$ a $\langle a', b', c', d' \rangle = 4\mathrm{ft}(\varphi', \psi', \mathbf{M})$
- $4\mathrm{ft}(\varphi, \psi, \mathbf{M})$ a $4\mathrm{ft}(\varphi', \psi', \mathbf{M})$ se týkají stejné matice dat, tedy vždy platí $a+b+c+d = a'+b'+c'+d'$
- K podmínce zachování pravdivosti $a' \leq a \wedge b' \leq b$ tedy můžeme přidat další požadavek $a+b+c+d = a'+b'+c'+d'$, který je pro $4\mathrm{ft}(\varphi, \psi, \mathbf{M})$ a $4\mathrm{ft}(\varphi', \psi', \mathbf{M})$ vždy splněn.
- Podmínku $a' \leq a \wedge b' \leq b \wedge a+b+c+d = a'+b'+c'+d'$ využijeme k definici slabě implikačních kvantifikátorů.

Slabě implikační kvantifikátory a pravidla - definice

- *4ft-kvantifikátor* $\approx$ *je slabě implikační* pokud pro všechny čtveřice $\langle a, b, c, d \rangle$ a $\langle a', b', c', d' \rangle$ celých nezáporných čísel platí:

  jestliže $\approx(a,b,c,d) = 1$ a zároveň $a' \geq a \wedge b' \leq b \wedge a+b+c+d = a'+b'+c'+d'$,
  pak i $\approx(a',b',c',d') = 1$.

- Podmínku $a' \geq a \wedge b' \leq b \wedge a+b+c+d = a'+b'+c'+d'$ nazýváme *podmínkou zachování pravdivosti pro implikační kvantifikátory*.

- *Asociační pravidlo* $\varphi \approx \psi$ *je slabě implikační* pokud 4ft-kvantifikátor $\approx$ je slabě implikační.

4ft-kvantifikátor $\rightarrow_{C,S}$ je slabě implikační

Je třeba ukázat:

Jestliže $\frac{a}{a+b} \geq C \wedge \frac{a}{a+b+c+d} \geq S$ a zároveň

$a' \geq a \wedge b' \leq b \wedge a+b+c+d = a'+b'+c'+d'$, pak i $\frac{a'}{a'+b'} \geq C \wedge \frac{a'}{a'+b'+c'+d'} \geq S$

- Z $a' \geq a \wedge b' \leq b$ postupně plyne: $\frac{a'}{a'+b'} \geq \frac{a}{a+b'} \geq \frac{a}{a+b} \geq C$

- Z $a' \geq a \wedge a+b+c+d = a'+b'+c'+d'$ postupně plyne:

$$\frac{a'}{a'+b'+c'+d'} = \frac{a'}{a+b+c+d} \geq \frac{a}{a+b+c+d} \geq S$$

Tedy $\rightarrow_{C,S}$ je slabě implikační.

Každý implikační kvantifikátor je slabě implikační

1. Předpokládáme, že $\approx$ je implikační, tedy že pro všechny čtveřice $\langle a, b, c, d \rangle$ a $\langle a', b', c', d' \rangle$ celých nezáporných čísel platí:

   jestliže $\approx(a,b,c,d) = 1$ a zároveň $a' \geq a$ a $b' \leq b$, pak i $\approx(a',b',c',d') = 1$.

2. Je třeba ukázat, že pro všechny čtveřice $\langle a, b, c, d \rangle$ a $\langle a', b', c', d' \rangle$ celých nezáporných čísel platí:

   jestliže $\approx(a,b,c,d) = 1$ a zároveň $a' \geq a \wedge b' \leq b \wedge a+b+c+d = a'+b'+c'+d'$, pak i $\approx(a',b',c',d') = 1$.

3. Nechť $\approx(a,b,c,d) = 1$. Z $a' \geq a \wedge b' \leq b \wedge a+b+c+d = a'+b'+c'+d'$ (tedy předpokladu pro 2) plyne $a' \geq a \wedge b' \leq b$ a podle bodu 1 platí i $\approx(a',b',c',d') = 1$. Tedy $\approx$ je slabě implikační.

Důležité slabě implikační kvantifikátory v proceduře 4ft-Miner (1)

- „Klasický“ kvantifikátor $\rightarrow_{C,S}(a,b,c,d) = 1$ právě když $\frac{a}{a+b} \geq C \wedge \frac{a}{a+b+c+d} \geq S$.

QUANTIFIERS	
PIM	p= 0.900
BASE	p= 50 Abs.

- Dolní kritická implikace se supportem: $\rightarrow^{!}_{p,\alpha,s}(a,b,c,d) = 1$ právě když

$$\sum_{i=a}^{a+b} \binom{a+b}{i} p^i (1-p)^{a+b-i} \leq \alpha \wedge \frac{a}{a+b+c+d} \geq s$$

QUANTIFIERS	
LCI	p= 0.900, Alpha= 0.050
BASE	p= 50 Abs.

Důležité slabě implikační kvantifikátory v proceduře 4ft-Miner (2)

- Horní kritická implikace $\rightarrow^{?}_{p,\alpha}(a,b,c,d) = 1$ právě když

$$\sum_{i=0}^{a} \binom{a+b}{i} p^i (1-p)^{a+b-i} > \alpha \ \wedge \frac{a}{a+b+c+d} \geq S$$

QUANTIFIERS	
LCI	p= 0.900, Alpha= 0.050
BASE	p= 50 Abs.