Tato prezentace je součástí wiki-prezentace Metoda GUHA a systém LISp-Miner
Je dostupná z této adresy
Verse 7. 9. 2019

Třídy 4ft-kvantifikátorů

Jan Rauch

Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická v Praze

Třídy 4ft-kvantifikátorů

- Podmínka zachování pravdivosti
- Třídy 4ft-kvantifikátorů
- Vztahy mezi třídami 4ft-kvantifikátorů

Podmínka zachování pravdivosti

- Třídy asociačních pravidel jsou definovány pomocí podmínek zachování pravdivosti.
- Podmínka zachování pravdivosti se týká dvou čtveřic nezáporných celých čísel $\langle a,b,c,d \rangle$ a $\langle a',b',c',d' \rangle$.
- Podmínka zachování pravdivosti (TPC Truth Preservation Condition) je {0,1}-hodnotová funkce $\mathbf{C}(a,b,c,d,a',b',c',d')$.

- Třída 4ft-kvantifikátorů 4ft[**C**] definovaná podmínkou zachování pravdivosti **C** je množina všech 4ft-kvantifikátorů ≈ splňujících podmínku:
  jestliže $\approx(a,b,c,d) = 1 \wedge \mathbf{C}(a,b,c,d,a',b',c',d') = 1$, pak $\approx(a',b',c',d') = 1$.

- Třída asociačních pravidel daná podmínkou zachování pravdivosti **C** je množina všech asociačních pravidel $\varphi \approx \psi$ takových, že ≈ patří do třídy 4ft-kvantifikátorů 4ft[**C**].

Třídy 4ft-kvantifikátorů (1)

Třída 4ft-kvantifikátorů	Podmínka zachování pravdivosti	Označení
implikační kvantifikátory	$a' \geq a \wedge b' \leq b$	IMP
slabě implikační kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a' \geq a \ \wedge\ b' \leq b$	WIMP
dvojitě implikační kvantifikátory	$a' \geq a \wedge b' \leq b \wedge c' \leq c$	DIMP
slabě dvojitě implikační kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a' \geq a \wedge b' \leq b \wedge c' \leq c$	WDIMP
Σ-dvojitě implikační kvantifikátory	$a' \geq a \wedge b'+c' \leq b+c$	ΣDIMP
slabě Σ-dvojitě implikační kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a' \geq a \wedge b'+c' \leq b+c$	WΣDIMP
ekvivalenční kvantifikátory	$a' \geq a \wedge b' \leq b \wedge c' \leq c \wedge d' \geq d$	EQ
slabě ekvivalenční kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a' \geq a \wedge b' \leq b \wedge c' \leq c \wedge d' \geq d$	WEQ
Σ-ekvivalenční kvantifikátory	$a'+d' \geq a+d \wedge b'+c' \leq b+c$	ΣEQ
slabě Σ-ekvivalenční kvantifikátory	$a+b+c+d = a'+b'+'c+d' \wedge$ $\wedge\ a'+d' \geq a+d$	WΣEQ
symetrické kvantifikátory	$a' = a \wedge b' = c \wedge c' = b \wedge d' = d$	SYM
s vlastností F	$(b \geq c-1 \geq 0 \wedge$ $\wedge\ a' = a \wedge b' = b+1 \wedge c' = c-1 \wedge d' = d) \vee$ $\vee\ (c \geq b-1 \geq 0 \wedge$ $\wedge\ a' = a \wedge b' = b-1 \wedge c' = c+1b \wedge d' = d)$	F-P

Třídy 4ft-kvantifikátorů (2)

Z hlediska dedukčních pravidel aplikovaných v souvislosti s procedurou 4ft-Miner jsou důležité zejména následující třídy které jsou podrobněji popsány:

- třída implikačních 4ft-kvantifikátorů
- třída slabě implikačních kvantifikátorů
- třída symetrických kvantifikátorů.

Vztahy mezi třídami 4ft-kvantifikátorů (1)

Základní tvrzení:
Nechť pro podmínky zachování pravdivosti $\mathbf{C_1}$ a $\mathbf{C_2}$ platí:
Jestliže $\mathbf{C}_2(a,b,c,d,a',b',c',d') = 1$, pak i $\mathbf{C}_1(a,b,c,d,a',b',c',d') = 1$, potom když 4ft-kvantifikátor ≈ patří do 4ft[$\mathbf{C}_1$] patří i do 4ft[$\mathbf{C}_2$].

Důkaz:

- Nechť ≈ patřící do 4ft[$\mathbf{C}_1$], máme dokázat že patří do 4ft[$\mathbf{C}_2$]. To znamená, že máme dokázat že když $\approx(a,b,c,d) = 1 \wedge \mathbf{C}_2(a,b,c,d,a',b',c',d') = 1$, pak $\approx(a',b',c',d') = 1$.

- Nechť tedy $\approx(a,b,c,d) = 1 \wedge \mathbf{C}_2(a,b,c,d,a',b',c',d') = 1$, pak ale platí i $\approx(a,b,c,d) = 1 \wedge \mathbf{C}_1(a,b,c,d,a',b',c',d') = 1$ protože předpokládáme, že když $\mathbf{C}_2(a,b,c,d,a',b',c',d') = 1$, pak i $\mathbf{C}_1(a,b,c,d,a',b',c',d') = 1$.

- Ale z $\approx(a,b,c,d) = 1 \wedge \mathbf{C}_1(a,b,c,d,a',b',c',d') = 1$ a z předpokladu, že ≈ patří do 4ft[$\mathbf{C}_1$] plyne $\approx(a',b',c',d') = 1$. Dokázali, že když $\approx(a,b,c,d) = 1 \wedge \mathbf{C}_2(a,b,c,d,a',b',c',d') = 1$, pak $\approx(a',b',c',d') = 1$. Tedy ≈ patří do 4ft[$\mathbf{C}_2$].

Vztahy mezi třídami 4ft-kvantifikátorů (2)

Příklady:

- $\mathbf{C}_{\mathbf{DIMP}}(a,b,c,d,a',b',c',d') = 1$ právě když $a' \geq a \wedge b' \leq b \wedge c' \leq c$
  $\mathbf{C}_{\mathbf{\Sigma DIMP}}(a,b,c,d,a',b',c',d') = 1$ právě když $a' \geq a \wedge b' + c' \leq b + c$
  Tedy když platí $\mathbf{C}_{\mathbf{DIMP}}$, tak platí $\mathbf{C}_{\mathbf{\Sigma DIMP}}$. To znamená, že každý Σ-dvojitě implikační kvantifikátor je také dvojitě implikační, symbolicky **ΣDIMP ⊆ DIMP**.

- $\mathbf{C}_{\mathbf{EQ}}(a,b,c,d,a',b',c',d') = 1$ právě když $a' \geq a \wedge b' \leq b \wedge c' \leq c \wedge d' \geq d$
  $\mathbf{C}_{\mathbf{\Sigma DIMP}}(a,b,c,d,a',b',c',d') = 1$ právě když $a' + d' \geq a + \mathrm{d} \wedge b' + c' \leq b + c$
  Tedy když platí $\mathbf{C}_{\mathbf{EQ}}$, tak platí $\mathbf{C}_{\mathbf{\Sigma EQ}}$. To znamená, že každý Σ-ekvivalenční kvantifikátor je také ekvivalenční, symbolicky **ΣEQ ⊆ EQ**.

Vztahy mezi třídami 4ft-kvantifikátorů (3)

Příklady:

- **C** $_{\mathbf{IMP}}(a,b,c,d,a',b',c',d')$ = 1 právě když $a' \geq a \wedge b' \leq b$
  C $_{\mathbf{WIMP}}(a,b,c,d,a',b',c',d')$ = 1 právě když $a+b+c+d = a'+b'+c'+d' \wedge a' \geq a \wedge b' \leq b$
  Tedy když platí **C** $_{\mathbf{WIMP}}$, tak platí **C** $_{\mathbf{IMP}}$. To znamená, že každý implikační kvantifikátor je také slabě implikační, symbolicky **IMP** $\subseteq$ **WIMP**.

Analogicky platí:

- Každý dvojitě implikační kvantifikátor je také slabě dvojitě implikační, symbolicky **DIMP** $\subseteq$ **WDIMP**.
- Každý Σ-dvojitě implikační kvantifikátor je také slabě Σ- dvojitě implikační, symbolicky **ΣDIMP** $\subseteq$ **WΣDIMP**.
- Každý ekvivalenční kvantifikátor je také slabě ekvivalenční, symbolicky **EQ** $\subseteq$ **WEQ**.
- Každý Σ- ekvivalenční kvantifikátor je také slabě Σ- ekvivalenční, symbolicky **ΣEQ** $\subseteq$ **WΣEQ**.

Vztahy mezi třídami 4ft-kvantifikátorů (4)

Krom výše definovaných tříd 4ft-kvantifikátorů je možno definovat řadu dalších tříd různým způsobem zajímavých. Zajímavé jsou i vztahy mezi nimi. Podrobnosti jsou v monografii Observational Calculi and Association Rules.

Zde se zajímáme o prakticky důležité třídy 4ft-kvantifikátorů implementovaných v proceduře 4ft-Miner. Praktická důležitost spočívá v možnosti aplikovat dedukční pravidla užitečná při implementaci procedury a při interpretaci jejich výsledků. Důležité jsou i možnosti využití vlastnosti tříd při zpracování neúplné informace.

Informace o takto prakticky důležitých třídách 4ft-kvantifikátorů jsou uvedeny dále. Poznamenejme, že možnosti využití dedukčních pravidel jsou důležité proto, že pracujeme s GUHA asociačními pravidly, která jsou obecnější než pravidla poskytovaná algoritmem apriori.

Je také třeba poznamenat, že užitečná dedukční pravidla existují pouze pro slabě implikační a symetrické kvantifikátory. Důkaz jejich korektnosti je snadný. Obtížnější je ukázat, že pro ostatní třídy neexistují vhodná dedukční pravidla. Část těchto důkazů je dostupná v článku Logical Aspects of Dealing with Domain Knowledge in Data Mining with Association Rules.