

Databázové systémy

Funkční závislosti

Vilém Vychodil

KMI/DATA1, Přednáška 11

Databázové systémy

Přednáška 11: Přehled

1 Funkční závislosti:

- pravdivost funkčních závislostí v systémech relací,
- teorie, modely,
- sémantické vyplývání,
- kanonické relace, kanonické modely,
- testování sémantického vyplývání,
- sémantický uzávěr množiny atributů,
- algoritmus pro výpočet uzávěru,
- charakterizace vyplývání pomocí sémantického uzávěru,
- vztah k pojmu klíč.

Funkční závislosti

motivace:

Formule $A \Rightarrow B$ ve tvaru implikace mezi konjunkcemi atributů vyjadřující podmínky ve tvaru když-pak – pokud jsou si hodnoty rovny na attributech z A , musí si být rovny na attributech z B , např. $\{\text{time, room}\} \Rightarrow \{\text{course}\}$.

související problémy:

- modely a sémantické vyplývání
- algoritmy pro rozhodování, jestli formule plyne z množiny formulí
- axiomatizace, automatické dokazovací systémy
- grafické dokazovací metody a normalizace důkazů
- minimální báze závislostí v databázových instancích
- vztah k pojmům nadklíč, klíč,...
- vztahy k fragmentům výrokových a predikátových logik
-

Syntax a sémantika funkčních závislostí

Definice (funkční závislost, angl.: *functional dependency*)

Nechť R je relační schéma. Pak **funkční závislost** nad schématem R je formule ve tvaru $A \Rightarrow B$, kde $A, B \subseteq R$.

Definice (pravdivost funkční závislosti v datech)

Nechť R je relační schéma a \mathcal{D} je relace nad schématem R . Pak funkční závislost $A \Rightarrow B$ nad schématem R je **pravdivá** v \mathcal{D} , což označujeme $\mathcal{D} \models A \Rightarrow B$, pokud pro každé n -tice $r_1, r_2 \in \mathcal{D}$ platí:

$$\text{pokud } r_1(A) = r_2(A), \text{ pak } r_1(B) = r_2(B).$$

V opačném případě říkáme, že $A \Rightarrow B$ neplatí v \mathcal{D} a píšeme $\mathcal{D} \not\models A \Rightarrow B$.

slovně: $\mathcal{D} \models A \Rightarrow B$ právě tehdy, když pro každé dvě n -tice r_1 a r_2 z \mathcal{D} platí, že pokud jsou si rovny na všech atributech z množiny A , pak jsou si rovny i na všech atributech z množiny B .

Příklad (Funkční závislosti, které jsou/nejsou pravdivé v \mathcal{D})

mějme relaci $\mathcal{D} = \{r_1, r_2, r_3, r_4\}$ nad schématem $R = \{\text{FOO}, \text{BAR}, \text{BAZ}, \text{QUX}\}$:

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333
10	22	a	444
20	33	a	555

$$r_1 = \{\langle \text{FOO}, 10 \rangle, \langle \text{BAR}, 22 \rangle, \langle \text{BAZ}, a \rangle, \langle \text{QUX}, 222 \rangle\}$$

$$r_2 = \{\langle \text{FOO}, 10 \rangle, \langle \text{BAR}, 33 \rangle, \langle \text{BAZ}, b \rangle, \langle \text{QUX}, 333 \rangle\}$$

$$r_3 = \{\langle \text{FOO}, 10 \rangle, \langle \text{BAR}, 22 \rangle, \langle \text{BAZ}, a \rangle, \langle \text{QUX}, 444 \rangle\}$$

$$r_4 = \{\langle \text{FOO}, 20 \rangle, \langle \text{BAR}, 33 \rangle, \langle \text{BAZ}, a \rangle, \langle \text{QUX}, 555 \rangle\}$$

$$\mathcal{D} \not\models \{\text{BAZ}\} \Rightarrow \{\text{BAR}\} \quad (\text{kvůli } r_1 \text{ a } r_4)$$

$$\mathcal{D} \not\models \{\text{FOO}, \text{BAZ}\} \Rightarrow \{\text{QUX}\} \quad (\text{kvůli } r_1 \text{ a } r_3)$$

$$\mathcal{D} \not\models \{\text{BAR}\} \Rightarrow \{\text{BAZ}\} \quad (\text{kvůli } r_2 \text{ a } r_4)$$

$$\mathcal{D} \models \{\text{FOO}, \text{BAR}\} \Rightarrow \{\text{BAZ}\}$$

$$\mathcal{D} \models \{\text{QUX}\} \Rightarrow S \quad (\text{pro jakékoliv } S \text{ je triviálně splněná})$$

$$\mathcal{D} \models \{\text{FOO}, \text{BAR}\} \Rightarrow \{\text{FOO}\}$$

Funkční závislosti jako integritní omezení

motivace:

Analogicky jako (nad)klíče lze chápat funkční závislosti jako integritní omezení omezující přípustné hodnoty relačních proměnných.

integritní omezení: $A \Rightarrow B$ musí být pravdivá v hodnotě r nad $R \supseteq A \cup B$:
pro $A = \{y_1, \dots, y_k\}$, $B = \{z_1, \dots, z_l\}$ uvažujeme

$$\sigma_{y_1=y'_1 \wedge \dots \wedge y_k=y'_k \wedge \neg(z_1=z'_1 \wedge \dots \wedge z_l=z'_l)}(\pi_{A \cup B}(r)) \bowtie \rho_{y'_1 \leftarrow y_1, \dots, y'_k \leftarrow y_k, z'_1 \leftarrow z_1, \dots, z'_l \leftarrow z_l}(\pi_{A \cup B}(r)) = \emptyset_{A \cup B},$$

předchozí výraz je pravdivý v \mathcal{D} pokud pro každé dvě n -tice $r_1, r_2 \in r^{\mathcal{D}}$ platí:

$$\text{pokud } r_1(A) = r_2(A), \text{ pak } r_1(B) = r_2(B)$$

Příklad (Tutorial D: Funkční závislosti jako integritní omezení)

```
VAR rt BASE
  RELATION {foo INT, bar INT, baz INT, qux INT}
  KEY {qux};

/* define constraint: {foo, bar} => {baz} */
CONSTRAINT fubar_fd
  IS_EMPTY ((rt TIMES (rt RENAME {PREFIX "" AS "n"}))
            WHERE AND {foo = nfoo, bar = nbar} AND
                    NOT (baz = nbaz));

/* define constraint: {foo} => {bar, baz} */
CONSTRAINT barbaz_fd
  IS_EMPTY ((rt TIMES (rt RENAME {PREFIX "" AS "n"}))
            WHERE AND (foo = nfoo) AND
                    NOT AND {bar = nbar, baz = nbaz});
```

Příklad (SQL: Funkční závislosti jako integritní omezení)

```
CREATE OR REPLACE FUNCTION fubar_fd_trg () RETURNS TRIGGER AS $$
BEGIN
    IF (0 < (SELECT count (*) FROM tab
            WHERE foo = NEW.foo AND bar = NEW.bar AND
            NOT (baz = NEW.baz))) THEN
        RAISE EXCEPTION 'FD_{foo,bar}=>_{baz}violated!';
    END IF;
    RETURN NEW;
END;
$$ LANGUAGE plpgsql;

CREATE TRIGGER fubar_fd BEFORE INSERT OR UPDATE ON tab
FOR EACH ROW EXECUTE PROCEDURE fubar_fd_trg ();
```


Triviální funkční závislosti

$A \Rightarrow B$ se nazývá **triviální**, pokud $\mathcal{D} \models A \Rightarrow B$ pro každou \mathcal{D} nad $R \supseteq A \cup B$

Věta (Charakterizace triviálních funkčních závislostí)

$A \Rightarrow B$ je triviální právě tehdy, když $B \subseteq A$.

Důkaz.

Pokud $B \not\subseteq A$, pak existuje $y \in B$ tak, že $y \notin A$. Uvažujme $\mathcal{D} = \{r_1, r_2\}$ nad $A \cup B$ tak, že $r_1(z) = 1$ pro každé $z \in A \cup B$ a

$$r_2(z) = \begin{cases} 1, & \text{pokud } z \in A, \\ 0, & \text{jinak.} \end{cases}$$

Pak zřejmě $\mathcal{D} \not\models A \Rightarrow B$, protože $r_1(A) = r_2(A)$, ale $r_1(y) \neq r_2(y)$ a $y \in B$.

Obráceně předpokládejme, že $B \subseteq A$. Pokud pro $r_1, r_2 \in \mathcal{D}$ platí, že $r_1(A) = r_2(A)$, tím spíš $r_1(B) = r_2(B)$ a tedy $\mathcal{D} \models A \Rightarrow B$. □

Věta (Vztah k pojmu funkce)

$\mathcal{D} \models A \Rightarrow B$ právě tehdy, když je $f = \{\langle r(A), r(B) \rangle \mid r \in \mathcal{D}\}$ zobrazení.

Důkaz.

Předpokládejme, že platí $\mathcal{D} \models A \Rightarrow B$ a uvažujme $\langle s, s_1 \rangle \in f$, $\langle s, s_2 \rangle \in f$. Stačí ukázat, že $s_1 = s_2$. Dle definice f existují $r_1, r_2 \in \mathcal{D}$ tak, že $r_1(A) = s = r_2(A)$. Odtud ale ihned užitím $\mathcal{D} \models A \Rightarrow B$ pro r_1 a r_2 dostáváme, že $r_1(B) = r_2(B)$, což znamená $s_1 = r_1(B) = r_2(B) = s_2$.

Obráceně, předpokládejme, že f je zobrazení a vezměme libovolné $r_1, r_2 \in \mathcal{D}$ takové, že $r_1(A) = r_2(A)$. Dle definice pravdivosti $A \Rightarrow B$ v \mathcal{D} stačí ukázat, že $r_1(B) = r_2(B)$. Dle definice f , $\langle r_1(A), r_1(B) \rangle \in f$ a $\langle r_2(A), r_2(B) \rangle \in f$. Jelikož je f zobrazení a $r_1(A) = r_2(A)$, musí platit $r_1(B) = r_2(B)$. \square

poznámka:

- $\mathcal{D} \models A \Rightarrow B$ p. k. f z věty je zobrazení z $\pi_A(\mathcal{D})$ do $\pi_B(\mathcal{D})$

Věta (O pravdivosti funkčních závislostí)

Pro každé $A, B, C \subseteq R$ a libovolné relace $\mathcal{D}, \mathcal{D}_1, \mathcal{D}_2$ nad R platí:

- 1 pokud $\mathcal{D} \models A \Rightarrow B$, pak $\mathcal{D} \models A \Rightarrow B \cap C$;
- 2 pokud $\mathcal{D} \models A \Rightarrow B$, pak $\mathcal{D} \models A \cup C \Rightarrow B$;
- 3 $\mathcal{D} \models A \Rightarrow B$ právě tehdy, když $\mathcal{D} \models A \Rightarrow B \setminus A$;
- 4 pokud $\mathcal{D}_1 \subseteq \mathcal{D}_2$ a $\mathcal{D}_2 \models A \Rightarrow B$, pak $\mathcal{D}_1 \models A \Rightarrow B$.

Důkaz.

Předpokládejme, že $\mathcal{D} \models A \Rightarrow B$ a vezměme libovolné $r_1, r_2 \in \mathcal{D}$ takové, že $r_1(A) = r_2(A)$. Dle předpokladu pak $r_1(B) = r_2(B)$ a tím spíš $r_1(B \cap C) = r_2(B \cap C)$, protože $B \cap C \subseteq B$, to dokazuje 1 a jednu stranu tvrzení 3. Analogicky 2. Pokud $\mathcal{D} \models A \Rightarrow B \setminus A$ a platí $r_1(A) = r_2(A)$, pak z $r_1(B \setminus A) = r_2(B \setminus A)$ a $r_1(A) = r_2(A)$ dostaneme $r_1(A \cup (B \setminus A)) = r_2(A \cup (B \setminus A))$, to jest platí 3, protože $B \subseteq A \cup (B \setminus A)$. Bod 4 je přímým důsledkem toho, že pro každé $r_1, r_2 \in \mathcal{D}_1$ platí $r_1, r_2 \in \mathcal{D}_2$. \square

Věta (Dekompozice na základě funkčních závislostí)

Mějme \mathcal{D} na R a $A, B \subseteq R$. Pokud $\mathcal{D} \models A \Rightarrow B$, pak \mathcal{D} má bezztrátovou dekompozici vzhledem k $A \cup B$ a $A \cup (R \setminus B)$.

Důkaz.

Máme ukázat, že $\pi_{A \cup B}(\mathcal{D}) \bowtie \pi_{A \cup (R \setminus B)}(\mathcal{D}) \subseteq \mathcal{D}$ (obrácená inkluze platí vždy).

Vezměme n -tice $r_1 \in \pi_{A \cup B}(\mathcal{D})$ a $r_2 \in \pi_{A \cup (R \setminus B)}(\mathcal{D})$, které jsou spojitelné. Jelikož

$$(A \cup B) \cap (A \cup (R \setminus B)) = A \cup (B \cap (R \setminus B)) = A \cup \emptyset = A,$$

platí, že $r_1(A) = r_2(A)$. Vezměme $s \in \mathcal{D}$ takovou, že $r_2 = s(A \cup (R \setminus B))$. Jelikož je B funkčně závislá na A , pak zřejmě z $r_1(A) = r_2(A) = s(A)$ dostáváme $r_1(B) = s(B)$, to jest $r_1 = s(A \cup B)$. To znamená, že $r_1 r_2 = s \in \mathcal{D}$. □

triviální použití:

- pokud je $A \Rightarrow B$ triviální, pak je $B \subseteq A$ a potom se \mathcal{D} na R dekomponuje na relace na schématech $A \cup B = A$ a $A \cup (R \setminus B) \supseteq A \cup (R \setminus A) = R$ (nezajímavé)

Příklad (Dekompozice použitím $\{DEPT\} \Rightarrow \{HEAD, SCHOOL, DEAN\}$)

SCHOOL	DEAN	DEPT	HEAD	ID	COURSE	YEAR
SCI	Blangis	AF	Durcet	7	QOPT1	2012
SCI	Blangis	AF	Durcet	8	LASR1	2012
SCI	Blangis	AF	Durcet	8	LASR1	2013
SCI	Blangis	CS	Curval	3	ALMA1	2012
SCI	Blangis	CS	Curval	3	ALMA1	2013
SCI	Blangis	CS	Curval	6	DATA1	2012
SCI	Blangis	CS	Curval	6	DATA1	2013
SCI	Blangis	CS	Curval	6	PAPR1	2012

=

SCHOOL	DEAN	DEPT	HEAD
SCI	Blangis	AF	Durcet
SCI	Blangis	CS	Curval

⊗

DEPT	ID	COURSE	YEAR
AF	7	QOPT1	2012
AF	8	LASR1	2012
AF	8	LASR1	2013
CS	3	ALMA1	2012
CS	3	ALMA1	2013
CS	6	DATA1	2012
CS	6	DATA1	2013
CS	6	PAPR1	2012

Pravdivost závislostí v systémech relací

Definice (pravdivost funkční závislosti v datech)

Nechť R je relační schéma a \mathcal{M} je množina relací nad schématem R . Položíme $\mathcal{M} \models A \Rightarrow B$ pokud $\mathcal{D} \models A \Rightarrow B$ pro každou $\mathcal{D} \in \mathcal{M}$.

Věta (pravdivost funkčních závislostí v systému dvouprvkových relací)

Pro každou \mathcal{D} nad R existuje konečný systém \mathcal{M} nejvýš dvouprvkových relací nad R tak, že pro každé $A, B \subseteq R$ platí: $\mathcal{D} \models A \Rightarrow B$ p. k. $\mathcal{M} \models A \Rightarrow B$.

Důkaz.

Položme $\mathcal{M} = \{\{r_1, r_2\} \mid r_1, r_2 \in \mathcal{D}\}$, to jest, \mathcal{M} je systém dvouprvkových tabulek složených ze všech n -tic z relace \mathcal{D} . Zřejmě platí, že $\mathcal{D} \models A \Rightarrow B$ p. k. pro libovolné $r_1, r_2 \in \mathcal{D}$ takové, že $r_1(A) = r_2(A)$, platí $r_1(B) = r_2(B)$, což je p. k. pro libovolné $r_1, r_2 \in \mathcal{D}$ platí, že $\{r_1, r_2\} \models A \Rightarrow B$ p. k. $\mathcal{M} \models A \Rightarrow B$. □

Příklad (Dvoupřvkové relace odpovídající \mathcal{D})

FOO	BAR	BAZ	QUX
10	22	a	222
20	33	a	555

FOO	BAR	BAZ	QUX
10	22	a	222
10	22	a	444

FOO	BAR	BAZ	QUX
10	22	a	444
10	33	b	333

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333
10	22	a	444
20	33	a	555

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333

FOO	BAR	BAZ	QUX
10	22	a	444
20	33	a	555

FOO	BAR	BAZ	QUX
10	33	b	333
20	33	a	555

Teorie a modely

motivace:

Chceme se zabývat tím, které funkční závislosti vyplývají z jiných závislostí. Primárně se zajímáme o sémantické vyplývání, pro jeho zavedení potřebujeme pojmy teorie a model.

Definice (teorie, angl.: *theory*)

Množinu funkčních závislostí (nad schématem R) nazveme **teorie** (nad R). Pokud je Γ teorie a $A \Rightarrow B \in \Gamma$, pak říkáme, že $A \Rightarrow B$ je **předpokladem** z Γ .

Definice (model, angl.: *model*)

Mějme teorii Γ . Relace \mathcal{D} je **model** Γ pokud pro každou $A \Rightarrow B \in \Gamma$ platí, že $\mathcal{D} \models A \Rightarrow B$. Množinu všech modelů Γ označujeme $\text{Mod}(\Gamma)$, to jest:

$$\text{Mod}(\Gamma) = \{\mathcal{D} \mid \text{pro každou } A \Rightarrow B \in \Gamma \text{ platí } \mathcal{D} \models A \Rightarrow B\}.$$

Příklad (Teorie a modely)

uvažujme následující teorii:

$$\Gamma = \{ \{ \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAR}, \text{BAZ} \}, \\ \{ \text{BAR}, \text{BAZ} \} \Rightarrow \{ \text{FOO} \}, \\ \{ \text{FOO}, \text{BAZ} \} \Rightarrow \{ \text{BAR} \}, \\ \{ \text{FOO}, \text{BAR} \} \Rightarrow \{ \text{BAZ} \} \}$$

potom pro danou Γ například:

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333
10	22	a	444
20	33	a	555

je model

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333
20	22	a	444
20	33	a	555

není model

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	222
10	22	a	444
20	33	a	555

není model

Sémantické vyplývání

Definice (sémantické vyplývání, angl.: *semantic entailment*)

Funkční závislost $A \Rightarrow B$ **sémanticky plyne** z teorie Γ pokud je $A \Rightarrow B$ pravdivá v každém modelu Γ , to znamená pokud $\text{Mod}(\Gamma) \models A \Rightarrow B$. Fakt, že $A \Rightarrow B$ sémanticky plyne z Γ značíme $\Gamma \models A \Rightarrow B$.

slovně:

Funkční závislost $A \Rightarrow B$ plyne z teorie Γ pokud je $A \Rightarrow B$ pravdivá v každé relaci, ve které jsou pravdivé všechny formule z teorie Γ .

speciální případ:

- pro $\Gamma = \emptyset$ píšeme $\models A \Rightarrow B$ místo $\emptyset \models A \Rightarrow B$
- význam: $\models A \Rightarrow B$ p. k. $A \Rightarrow B$ je pravdivá v každé relaci
- z předchozího víme: $\models A \Rightarrow B$ p. k. $A \Rightarrow B$ je triviální, to jest $B \subseteq A$

Věta (Vlastnosti teorií, modelů a sémantického vyplývání)

Pro libovolné teorie $\Gamma, \Gamma_1, \Gamma_2$ nad R platí:

- 1 pokud $\Gamma_1 \subseteq \Gamma_2$, pak $\text{Mod}(\Gamma_2) \subseteq \text{Mod}(\Gamma_1)$;
- 2 $\text{Mod}(\emptyset)$ je množina všech relací nad R ;
- 3 $\text{Mod}(\Gamma) \neq \emptyset$ pro každou teorii Γ ;
- 4 pokud $\Gamma_1 \subseteq \Gamma_2$ a $\Gamma_1 \models A \Rightarrow B$, pak $\Gamma_2 \models A \Rightarrow B$;
- 5 $\Gamma \models A \Rightarrow B$ pro každou $A \Rightarrow B \in \Gamma$.

Důkaz.

Pokud $\mathcal{D} \in \text{Mod}(\Gamma_2)$, pak pro každou $A \Rightarrow B \in \Gamma_2$ platí, že $\mathcal{D} \models A \Rightarrow B$. Jelikož ale $\Gamma_1 \subseteq \Gamma_2$, tím spíše platí $\mathcal{D} \models A \Rightarrow B$ pro každou $A \Rightarrow B \in \Gamma_1$, což ukazuje 1; 2 plyne přímo z definice; 3 platí, protože $\mathcal{D} \in \text{Mod}(\Gamma)$ pro každou \mathcal{D} splňující $|\mathcal{D}| < 2$; 4 je důsledkem 1; 5 platí triviálně, protože $\mathcal{D} \models A \Rightarrow B$ pro každé $A \Rightarrow B \in \Gamma$ a $\mathcal{D} \in \text{Mod}(\Gamma)$. □

Příklad (Sémantické důsledky teorie)

teorie z předchozího příkladu:

$$\Gamma = \{ \{ \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAR}, \text{BAZ} \}, \{ \text{BAR}, \text{BAZ} \} \Rightarrow \{ \text{FOO} \}, \\ \{ \text{FOO}, \text{BAZ} \} \Rightarrow \{ \text{BAR} \}, \{ \text{FOO}, \text{BAR} \} \Rightarrow \{ \text{BAZ} \} \}$$

netriviální funkční závislosti $A \Rightarrow B$, kde $A \cap B = \emptyset$ a $\Gamma \models A \Rightarrow B$:

$\{ \text{BAR}, \text{BAZ}, \text{QUX} \} \Rightarrow \{ \text{FOO} \},$	$\{ \text{BAR}, \text{BAZ} \} \Rightarrow \{ \text{FOO} \},$	$\{ \text{BAR}, \text{QUX} \} \Rightarrow \{ \text{BAZ} \},$
$\{ \text{BAR}, \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAZ} \},$	$\{ \text{BAR}, \text{QUX} \} \Rightarrow \{ \text{FOO} \},$	$\{ \text{BAZ}, \text{QUX} \} \Rightarrow \{ \text{BAR} \},$
$\{ \text{BAZ}, \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAR} \},$	$\{ \text{BAZ}, \text{QUX} \} \Rightarrow \{ \text{FOO} \},$	$\{ \text{FOO}, \text{BAR}, \text{QUX} \} \Rightarrow \{ \text{BAZ} \},$
$\{ \text{FOO}, \text{BAR} \} \Rightarrow \{ \text{BAZ} \},$	$\{ \text{FOO}, \text{BAZ}, \text{QUX} \} \Rightarrow \{ \text{BAR} \},$	$\{ \text{FOO}, \text{BAZ} \} \Rightarrow \{ \text{BAR} \},$
$\{ \text{FOO}, \text{QUX} \} \Rightarrow \{ \text{BAR}, \text{BAZ} \},$	$\{ \text{FOO}, \text{QUX} \} \Rightarrow \{ \text{BAR} \},$	$\{ \text{FOO}, \text{QUX} \} \Rightarrow \{ \text{BAZ} \},$
$\{ \text{QUX} \} \Rightarrow \{ \text{BAR}, \text{BAZ} \},$	$\{ \text{QUX} \} \Rightarrow \{ \text{BAR} \},$	$\{ \text{QUX} \} \Rightarrow \{ \text{BAZ} \},$
$\{ \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAR}, \text{BAZ} \},$	$\{ \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAR} \},$	$\{ \text{QUX} \} \Rightarrow \{ \text{FOO}, \text{BAZ} \},$
$\{ \text{QUX} \} \Rightarrow \{ \text{FOO} \}$		

Kanonické relace a modely

Definice (kanonická relace, angl.: *canonical relation*)

Pro libovolnou $M \subseteq R$ definujeme relaci \mathcal{D}_M nad schématem R tak, že $\mathcal{D}_M = \{r_1, r_2\}$, přitom $r_1(y) = p$ pro každý $y \in R$ a

$$r_2(y) = \begin{cases} p, & \text{pokud } y \in M, \\ q, & \text{pokud } y \notin M, \end{cases}$$

přitom p, q označují dva různé (fixní) prvky domén atributů z R . Takto zavedenou relaci \mathcal{D}_M nazveme **kanonická relace** nad schématem R .

Definice (kanonický model, angl.: *canonical model*)

Mějme teorii Γ . Kanonická relace \mathcal{D}_M , která je modelem Γ , se nazývá **kanonický model** Γ . Množinu všech kanonických modelů Γ značíme $\text{Mod}_C(\Gamma)$, to jest:

$$\text{Mod}_C(\Gamma) = \{\mathcal{D}_M \mid \mathcal{D}_M \in \text{Mod}(\Gamma)\}.$$

Věta (Pravdivost funkčních závislostí v kanonických relacích)

Pro libovolné $A, B, M \subseteq R$ jsou následující tvrzení ekvivalentní:

- 1 $\mathcal{D}_M \models A \Rightarrow B$,
- 2 pokud $A \subseteq M$, pak $B \subseteq M$.

Důkaz.

Nejprve ukážeme, že pro $\mathcal{D}_M = \{r_1, r_2\}$ a libovolnou $C \subseteq R$ platí, že $r_1(C) = r_2(C)$ p. k. $C \subseteq M$. Dle definice kanonické relace máme, že $r_1(C) = r_2(C)$ znamená $r_2(y) = p$ pro každý $y \in C$. To znamená, $y \in M$ pro každý $y \in C$, to jest $C \subseteq M$.

S použitím tohoto faktu, $\mathcal{D}_M \models A \Rightarrow B$ právě tehdy, když $r_1(A) = r_2(A)$ implikuje $r_1(B) = r_2(B)$, to platí právě tehdy, když $A \subseteq M$ implikuje $B \subseteq M$. To jest body 1 a 2 jsou ekvivalentní. □

poznámka:

- pozorování slouží ke zjednodušenému vyjádření $\mathcal{D}_M \models A \Rightarrow B$

Věta (Charakterizace \models pomocí kanonických modelů)

$\Gamma \models A \Rightarrow B$ právě tehdy, když $\text{Mod}_C(\Gamma) \models A \Rightarrow B$.

Důkaz.

Pokud $\Gamma \models A \Rightarrow B$, to jest dle definice $\text{Mod}(\Gamma) \models A \Rightarrow B$ a tím spíš $\text{Mod}_C(\Gamma) \models A \Rightarrow B$, protože $\text{Mod}_C(\Gamma) \subseteq \text{Mod}(\Gamma)$.

Obráceně, předpokládejme, že $\text{Mod}_C(\Gamma) \models A \Rightarrow B$ a vezměme libovolný $\mathcal{D} \in \text{Mod}(\Gamma)$. To znamená, že $\mathcal{D} \models C \Rightarrow D$ pro každou $C \Rightarrow D \in \Gamma$. Dle jednoho z předchozích tvrzení existuje množina dvouprvkových relací \mathcal{M} tak, že $\mathcal{M} \models C \Rightarrow D$ pro každou $C \Rightarrow D \in \Gamma$. Ke každé $\mathcal{D}' \in \mathcal{M}$ můžeme vzít množinu atributů, na kterých jsou si obě n -nice z \mathcal{D}' rovny, to jest

$$M = \{y \in R \mid r_1(y) = r_2(y)\},$$

kde $\mathcal{D}' = \{r_1, r_2\}$. Potom \mathcal{D}_M je kanonický model Γ . Jelikož $\text{Mod}_C(\Gamma) \models A \Rightarrow B$, pak i $\mathcal{D}_M \models A \Rightarrow B$. Protože jsme vzali $\mathcal{D}' \in \mathcal{M}$ libovolně, máme $\mathcal{D} \models A \Rightarrow B$. \square

Příklad (Kanonické relace odpovídající \mathcal{D})

FOO	BAR	BAZ	QUX
10	22	a	222
20	33	a	555

FOO	BAR	BAZ	QUX
10	22	a	222
10	22	a	444

FOO	BAR	BAZ	QUX
10	22	a	444
10	33	b	333

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333
10	22	a	444
20	33	a	555

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333

FOO	BAR	BAZ	QUX
10	22	a	444
20	33	a	555

FOO	BAR	BAZ	QUX
10	33	b	333
20	33	a	555

Příklad (Kanonické relace odpovídající \mathcal{D})

FOO	BAR	BAZ	QUX
0	0	1	0
1	1	1	1

FOO	BAR	BAZ	QUX
1	1	1	0
1	1	1	1

FOO	BAR	BAZ	QUX
1	0	0	0
1	1	1	1

FOO	BAR	BAZ	QUX
10	22	a	222
10	33	b	333
10	22	a	444
20	33	a	555

FOO	BAR	BAZ	QUX
1	0	0	0
1	1	1	1

FOO	BAR	BAZ	QUX
0	0	1	0
1	1	1	1

FOO	BAR	BAZ	QUX
0	1	0	0
1	1	1	1

Vztah k sémantickému vyplývání z výrokové logiky

motivace:

Funkční závislosti lze chápat jako výrokové formule. V jakém vztahu je tedy výrokové sémantické vyplývání a \models tak, jak jsme jej zavedli my?

funkční závislosti jako výrokové formule:

$$\{y_1, \dots, y_m\} \Rightarrow \{z_1, \dots, z_n\} \text{ vs. } (y_1 \wedge \dots \wedge y_m) \Rightarrow (z_1 \wedge \dots \wedge z_n)$$

Pro $A \Rightarrow B$, Γ a \mathcal{D}_M lze proto brát jejich výrokové protějšky:

- odpovídající výrokovou formuli $\varphi_{A,B}$;
 - odpovídající výrokovou teorii $\Gamma_V = \{\varphi_{A,B} \mid A \Rightarrow B \in \Gamma\}$;
 - odpovídající ohodnocení výrokových proměnných: $e_M(y) = 1$ p. k. $y \in M$;
- přitom jako důsledek předchozí věty máme:

$$\Gamma \models A \Rightarrow B \text{ p. k. } e(\varphi_{A,B}) = 1 \text{ pro každý model } e \text{ teorie } \Gamma_V.$$

důsledek: vyplývání funkčních závislostí lze ověřovat *tabulkovou metodou*

Příklad (Test sémantického vyplývání pomocí tabelace)

$M \subseteq R$	φ_1	φ_2	φ_3	φ_4	ψ_1	ψ_2	ψ_3
$\{\}$	1	1	1	1	1	1	1
$\{\text{QUX}\}$	0	1	1	1	1	1	1
$\{\text{BAZ}\}$	1	1	1	1	1	0	1
$\{\text{BAZ}, \text{QUX}\}$	0	1	1	1	1	0	0
$\{\text{BAR}\}$	1	1	1	1	1	1	1
$\{\text{BAR}, \text{QUX}\}$	0	1	1	1	1	1	1
$\{\text{BAR}, \text{BAZ}\}$	1	0	1	1	1	0	1
$\{\text{BAR}, \text{BAZ}, \text{QUX}\}$	0	0	1	1	1	0	1
$\{\text{FOO}\}$	1	1	1	1	1	1	1
$\{\text{FOO}, \text{QUX}\}$	0	1	1	1	1	1	1
$\{\text{FOO}, \text{BAZ}\}$	1	1	0	1	0	0	1
$\{\text{FOO}, \text{BAZ}, \text{QUX}\}$	0	1	0	1	1	0	0
$\{\text{FOO}, \text{BAR}\}$	1	1	1	0	1	1	1
$\{\text{FOO}, \text{BAR}, \text{QUX}\}$	0	1	1	0	1	1	1
$\{\text{FOO}, \text{BAR}, \text{BAZ}\}$	1	1	1	1	0	1	1
$\{\text{FOO}, \text{BAR}, \text{BAZ}, \text{QUX}\}$	1	1	1	1	1	1	1

teorie:

$$\varphi_1: \{\text{QUX}\} \Rightarrow \{\text{FOO}, \text{BAR}, \text{BAZ}\}$$

$$\varphi_2: \{\text{BAR}, \text{BAZ}\} \Rightarrow \{\text{FOO}\},$$

$$\varphi_3: \{\text{FOO}, \text{BAZ}\} \Rightarrow \{\text{BAR}\},$$

$$\varphi_4: \{\text{FOO}, \text{BAR}\} \Rightarrow \{\text{BAZ}\}$$

testované formule:

$$\psi_1: \{\text{FOO}, \text{BAZ}\} \Rightarrow \{\text{QUX}\}$$

$$\psi_2: \{\text{BAZ}\} \Rightarrow \{\text{FOO}, \text{BAR}\}$$

$$\psi_3: \{\text{BAZ}, \text{QUX}\} \Rightarrow \{\text{BAR}\}$$

algoritmus pro výpočet $[M]_{\Gamma}$

Data: $M \subseteq R$ a teorie Γ nad R

Result: množina $[M]_{\Gamma} \subseteq R$ (**sémantický uzávěr M vzhledem k Γ**)

$W := M$; /* W je pomocná proměnná */

repeat

$L := W$; /* L označuje poslední vypočtenou hodnotu W */

foreach $E \Rightarrow F \in \Gamma$ **do**

if $E \subseteq W$ **then**

$W := W \cup F$; /* aktualizuj W */

$\Gamma := \Gamma \setminus \{E \Rightarrow F\}$; /* $E \Rightarrow F$ už není potřeba v Γ */

end

end

until $L = W$;

return W ;

proč zavádíme: pomocí $[M]_{\Gamma}$ ukážeme efektivní test sémantického vyplývání

Příklad (Průběh výpočtu sémantického uzávěru)

$W:$ $\Gamma:$
 $\{A\}$ $\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}$
 $\{A, B, C\}$

$W:$ $\Gamma:$
 $\{A, E\}$ $\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}$
 $\{A, B, C, E\}$ $\{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}$
 $\{A, B, C, E, G\}$

$W:$ $\Gamma:$
 $\{E, F\}$ $\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}$
 $\{A, B, E, F\}$ $\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{D, G\} \Rightarrow \{A, C, H\}$
 $\{A, B, C, E, F\}$ $\{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{D, G\} \Rightarrow \{A, C, H\}$
 $\{A, B, C, E, F, G\}$

odtud: $[\{A\}]_{\Gamma} = \{A, B, C\}, [\{A, E\}]_{\Gamma} = \{A, B, C, E, G\}, [\{E, F\}]_{\Gamma} = \{A, B, C, E, F, G\}$

Věta (Základní vlastnost sémantického uzávěru I)

Pro libovolnou Γ nad R a $A \subseteq R$ platí, že $\Gamma \models A \Rightarrow [A]_{\Gamma}$.

Důkaz.

Nechť \mathcal{D} je libovolný model $\mathcal{D} \in \text{Mod}(\Gamma)$. Vezměme libovolné $r_1, r_2 \in \mathcal{D}$ takové, že $r_1(A) = r_2(A)$. Stačí ukázat, že $r_1([A]_{\Gamma}) = r_2([A]_{\Gamma})$. Toto tvrzení prokážeme indukcí přes počet kroků algoritmu pro výpočet $[A]_{\Gamma}$.

Pro počáteční hodnotu $W = A$ je tvrzení zřejmé, protože $A \Rightarrow A$ je triviální funkční závislost. Předpokládejme, že pro W platí $r_1(W) = r_2(W)$ a uvažujme $E \Rightarrow F \in \Gamma$ takovou, že $E \subseteq W$. Pak stačí ukázat, že $r_1(W \cup F) = r_2(W \cup F)$. Z faktů, že $\mathcal{D} \in \text{Mod}(\Gamma)$, $r_1, r_2 \in \mathcal{D}$, $r_1(W) = r_2(W)$, $E \subseteq W$ a $E \Rightarrow F \in \Gamma$ ihned dostáváme, že $r_1(F) = r_2(F)$ což dohromady s naším předpokladem $r_1(W) = r_2(W)$ dává, že $r_1(W \cup F) = r_2(W \cup F)$.

Odtud plyne, že pro každou průběžnou hodnotu W v algoritmu je $\mathcal{D} \models A \Rightarrow W$, tedy $\Gamma \models A \Rightarrow W$. Zbytek plyne z toho, že $[A]_{\Gamma}$ je poslední hodnotou W . □

Věta (Základní vlastnost sémantického uzávěru II)

Pokud $\Gamma \models A \Rightarrow B$, pak $B \subseteq [A]_{\Gamma}$.

Důkaz.

Tvrzení prokážeme obměnou: Za předpokladu, že $B \not\subseteq [A]_{\Gamma}$ ukážeme, že $\Gamma \not\models A \Rightarrow B$. K tomu stačí ukázat, že za předpokladu $B \not\subseteq [A]_{\Gamma}$ existuje model $\mathcal{D} \in \text{Mod}(\Gamma)$ tak, že $\mathcal{D} \not\models A \Rightarrow B$.

Předpokládejme tedy, že $B \not\subseteq [A]_{\Gamma}$. Hledaný model budeme uvažovat jako kanonickou relaci. Konkrétně vezmeme $\mathcal{D}_{[A]_{\Gamma}}$. Z věty o platnosti $A \Rightarrow B$ v kanonické relaci dostáváme, že $\mathcal{D}_{[A]_{\Gamma}} \not\models A \Rightarrow B$, protože $A \subseteq [A]_{\Gamma}$, ale $B \not\subseteq [A]_{\Gamma}$ (předpoklad). Stačí tedy ukázat, že $\mathcal{D}_{[A]_{\Gamma}} \in \text{Mod}(\Gamma)$.

Vezměme libovolnou $E \Rightarrow F \in \Gamma$. Z algoritmu pro výpočet $[A]_{\Gamma}$ je patrné, že pokud $E \subseteq [A]_{\Gamma}$, pak musí platit i $F \subseteq [A]_{\Gamma}$, protože $[A]_{\Gamma}$ je poslední průběžnou hodnotou W z algoritmu. To jest, $E \subseteq [A]_{\Gamma}$ implikuje $F \subseteq [A]_{\Gamma}$, což dává $\mathcal{D}_{[A]_{\Gamma}} \models E \Rightarrow F$ a tedy $\mathcal{D}_{[A]_{\Gamma}} \in \text{Mod}(\Gamma)$, protože $E \Rightarrow F \in \Gamma$ jsme zvolili libovolně. \square

Věta (Charakterizace sémantického vyplývání)

Pro každou teorii Γ a $A, B \subseteq R$ jsou následující tvrzení ekvivalentní:

- 1 $\Gamma \models A \Rightarrow B$,
- 2 $B \subseteq [A]_{\Gamma}$,
- 3 $\mathcal{D}_{[A]_{\Gamma}} \models A \Rightarrow B$.

Důkaz.

Z bodu 1 plyne bod 2 díky předchozí větě. Předpokládejme, že platí 2, pak ihned dostáváme 3 z věty o platnosti funkčních závislostí v kanonických relacích. Předpokládejme, že platí 3. Pak z $A \subseteq [A]_{\Gamma}$ ihned dostáváme, že $B \subseteq [A]_{\Gamma}$. Jelikož ale $\Gamma \models A \Rightarrow [A]_{\Gamma}$ (viz základní vlastnost sémantického uzávěru), tím spíš i $\Gamma \models A \Rightarrow B$, protože $B \subseteq [A]_{\Gamma}$. □

důsledky:

- $[A]_{\Gamma}$ je největší prvek množiny $\{B \subseteq R \mid \Gamma \models A \Rightarrow B\}$
- podmínka 2 je *efektivní test sémantického vyplývání*

Příklad (Testování sémantického vyplývání pomocí uzávěrů)

Pro $\Gamma = \{\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}\}$ z předchozího příkladu a množinu atributů $M = \{A, E\}$ máme:

$W:$	$\Gamma:$
$\{A, E\}$	$\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}$
$\{A, B, C, E\}$	$\{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}$
$\{A, B, C, E, G\}$	

To jest $[M]_{\Gamma} = \{A, B, C, E, G\}$ a platí, že $\Gamma \models M \Rightarrow N$ p. k. $N \subseteq \{A, B, C, E, G\}$.

Například tedy platí:

$\Gamma \models \{A, E\} \Rightarrow \{G\}$	$\Gamma \models \{A, E\} \Rightarrow \{C\}$	$\Gamma \models \{A, E\} \Rightarrow \{C, G\}$
$\Gamma \models \{A, E\} \Rightarrow \{B, G\}$	$\Gamma \models \{A, E\} \Rightarrow \{B, C\}$	$\Gamma \models \{A, E\} \Rightarrow \{B, C, G\}$

Na druhou stranu například:

$\Gamma \not\models \{A, E\} \Rightarrow \{D\}$	$\Gamma \not\models \{A, E\} \Rightarrow \{H\}$	$\Gamma \not\models \{A, E\} \Rightarrow \{D, E\}$
$\Gamma \not\models \{A, E\} \Rightarrow \{F, G\}$	$\Gamma \not\models \{A, E\} \Rightarrow \{C, D\}$	$\Gamma \not\models \{A, E\} \Rightarrow \{C, G, H\}$

Klíče z pohledu funkčních závislostí

Definice (**nadklíč**, **klíč**, angl.: *superkey*, *key*)

Mějme relační schéma R a teorii Γ nad R . Pak **nadklíč** schématu R vzhledem k Γ je libovolná $K \subseteq R$ taková, že $\Gamma \models K \Rightarrow R$. Pokud je K nadklíč R vzhledem k Γ a žádná $K' \subset K$ není nadklíč R vzhledem k Γ , pak je K **klíč** R vzhledem k Γ .

otázka:

Jak souvisí s pojmem klíče relační proměnné (PŘEDNÁŠKA 2)?

následující je ekvivalentní:

- 1 K_1, \dots, K_n je množina klíčů (relační proměnné) typu R ;
to jest $\{K_1, \dots, K_n\} \neq \emptyset$ a $K_i \not\subseteq K_j$ pro každé $i \neq j$ (PŘEDNÁŠKA 2) p. k.
- 2 každý K_i je klíč R vzhledem k $\Gamma = \{K_1 \Rightarrow R, \dots, K_n \Rightarrow R\}$.

pozorování: každý nadklíč (R vzhledem k Γ) obsahuje nějaký klíč (R vzhledem k Γ)

Příklad (Nadklíče a klíče)

Pro $\Gamma = \{\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}\}$ z předchozích příkladů jsou klíče schématu $R = \{A, \dots, H\}$ vzhledem k Γ následující:

$$K_1 = \{A, D\}, \quad K_2 = \{B, D\}, \quad K_3 = \{C, D, E\}, \quad K_4 = \{D, F\}, \quad K_5 = \{D, G\},$$

protože pro K_1 platí:

$$[\{A, D\}]_{\Gamma} = R, \quad [\{A\}]_{\Gamma} = \{A, B, C\} \neq R, \quad [\{D\}]_{\Gamma} = \{D\} \neq R,$$

to znamená: $\Gamma \models K_1 \Rightarrow R$ a $\Gamma \not\models M \Rightarrow R$ pro každou $M \subset K_1$, to jest K_1 je klíč. Analogicky se dá ukázat pro K_2, \dots, K_5 . Například pro K_3 je $[K_3]_{\Gamma} = R$ a

$$[\{C, D\}]_{\Gamma} = \{C, D\} \neq R, \quad [\{C, E\}]_{\Gamma} = \{C, E, G\} \neq R, \quad [\{D, E\}]_{\Gamma} = \{D, E\} \neq R.$$

Následující jsou nadklíče, ale nejsou klíče:

$$\{A, D, E\}, \quad \{B, D, G\}, \quad \{C, D, E, F\}, \quad \{D, F, H\}, \quad \{A, \dots, H\}$$

Například $\{A\}$, $\{A, E\}$, $\{E, F\}$ nejsou nadklíče, protože $[\{A\}]_{\Gamma} = \{A, B, C\} \neq R$, $[\{A, E\}]_{\Gamma} = \{A, B, C, E, G\} \neq R$ a $[\{E, F\}]_{\Gamma} = \{A, B, C, E, F, G\} \neq R$.

Příklad (Nalezení klíče postupnou redukcí nadklíče)

Pokud je dána Γ nad R , (některý) klíč schématu R vzhledem k Γ lze nalézt tak, že vyjdeme z nadklíče $K = R$ a postupně z něj odebíráme atributy, dokud je množina pořád nadklíč. Pokud už žádný atribut nelze odebrat, výsledná množina K je klíč.

Pro $\Gamma = \{\{A\} \Rightarrow \{B, C\}, \{B, D\} \Rightarrow \{E, F\}, \{C, E\} \Rightarrow \{G\}, \{F\} \Rightarrow \{A, B\}, \{D, G\} \Rightarrow \{A, C, H\}\}$ z předchozích příkladů můžeme najít klíče K_1, K_2, K_3 takto:

$\{A, B, C, D, E, F, G, H\}$

$\{A, B, C, D, E, G, H\}$

$\{A, B, D, E, G, H\}$

$\{A, B, D, E, H\}$

$\{A, D, E, H\}$

$\{A, D, E\}$

$\{A, D\} = K_1$

$\{A, B, C, D, E, F, G, H\}$

$\{A, B, C, D, E, F, H\}$

$\{A, B, D, E, F, H\}$

$\{A, B, D, F, H\}$

$\{A, B, D, H\}$

$\{A, B, D\}$

$\{B, D\} = K_2$

$\{A, B, C, D, E, F, G, H\}$

$\{A, B, C, D, E, G, H\}$

$\{A, B, C, D, E, H\}$

$\{B, C, D, E, H\}$

$\{B, C, D, E\}$

$\{C, D, E\} = K_3$

Analogicky pro K_4 a K_5 .




pozor: algoritmus závisí na výběru prvku z aktuálního nadklíče (!!!)

Přednáška 11: Závěr

pojmy k zapamatování:

- syntax a sémantika funkčních závislostí
- teorie, model, sémantické vyplývání
- kanonická relace, kanonický model
- sémantický uzávěr množiny atributů, charakterizace sémantického vyplývání

použité zdroje:

-  Date C. J.: *Database in Depth: Relational Theory for Practitioners*
O'Reilly Media 2005, ISBN 978-0596100124
-  Maier D: *Theory of Relational Databases*
Computer Science Press 1983, ISBN 978-0914894421
-  Simovici D.: Tenney R.: *Relational Database Systems*
Academic Press 1995, ISBN 978-0126443752