Katedra informatiky
Přírodovědecká fakulta
Univerzita Palackého v Olomouci

DIPLOMOVÁ PRÁCE

AD formule v podmíněné formální konceptuální analýze

2024

Jan Holcman

Vedoucí práce:
doc. RNDr. Michal Krupka, Ph.D.

Studijní program: Informatika,
Specializace: Obecná informatika

Bibliografické údaje

Autor:	Jan Holcman
Název práce:	AD formule v podmíněné formální konceptuální analýze
Typ práce:	diplomová práce
Pracoviště:	Katedra informatiky, Přírodovědecká fakulta, Univerzita Palackého v Olomouci
Rok obhajoby:	2024
Studijní program:	Informatika, Specializace: Obecná informatika
Vedoucí práce:	doc. RNDr. Michal Krupka, Ph.D.
Počet stran:	34
Přílohy:	elektronická data v úložišti katedry informatiky
Jazyk práce:	český

Bibliographic info

Author:	Jan Holcman
Title:	
Thesis type:	master thesis
Department:	Department of Computer Science, Faculty of Science, Palacký University Olomouc
Year of defense:	2024
Study program:	Computer Science, Specialization: General Computer Science
Supervisor:	doc. RNDr. Michal Krupka, Ph.D.
Page count:	34
Supplements:	electronic data in the storage of department of computer science
Thesis language:	Czech

Anotace

AD formule jsou užitečným nástrojem při analýze binárních dat metodou formální konceptuální analýzy (FCA). Podmíněná FCA je potom varianta FCA, kterou lze použít při neúplnosti vstupních dat. Tato práce poskytuje úvod jak do teorie AD formulí, tak i úvod do teorie podmíněné FCA, a jsou zde prezentovány nové výsledky související s implementací AD formulí v podmíněné FCA.

Synopsis

AD formulas are a useful tool when conducting the analysis of binary data through the method of formal concept analysis (FCA). Conditional FCA is a variant of FCA that can be used when the input data are incomplete. This work provides an introduction to both the theory of AD formulas and the theory of conditional FCA, and presents new results related to the implementation of AD formulas in conditional FCA.

Klíčová slova: Formální konceptuální analýza (FCA); AD formule; neúplná informace; podmíněná množina; podmíněná FCA

Keywords: Formal concept analysis (FCA); AD formula; incomplete information; conditional set; conditional FCA

Děkuji doc. RNDr. Michalovi Krupkovi, PhD. za nemalé množství času, které investoval do mého vzdělání, za vedení této práce a cenné rady. Dále děkuji všem mým blízkým za vytrvalou podporu při studiu.

Odevzdáním tohoto textu jeho autor/ka místopřísežně prohlašuje, že celou práci včetně příloh vypracoval/a samostatně a za použití pouze zdrojů citovaných v textu práce a uvedených v seznamu literatury.

Obsah

1 Úvod

Formální konceptuální analýza, zkráceně FCA, je metoda analýzy tabulkových binárních dat. Taková data lze přirozeně interpretovat jako vztah mezi objekty, kterým odpovídají řádky tabulky, a atributy, kterým odpovídají sloupce. Jednotlivé hodnoty v tabulce, řekněme jedničky a nuly, potom rozlišují, zda objekt na příslušném řádku má (hodnota 1) nebo nemá (hodnota 0) atribut v příslušném sloupci.

Při analýze jsou (mimo jiné) identifikovány určité shluky jedniček, nazývané formální koncepty, kterými lze reprezentovat pojmy obsažené v datech. Společně s vhodným uspořádáním tvoří množina všech formálních konceptů úplný svaz, ten nazýváme konceptuální svaz a je jedním z hlavních výstupů analýzy. Základy FCA jsou popsány v kapitole 2.

Je známo, že konceptuální svaz může obsahovat až exponenciálně mnoho formálních konceptů vzhledem k velikosti vstupních dat. Tato skutečnost byla jednou z motivací ke vzniku takzvaných AD formulí [3]. Jedná se o doplňkový nástroj k FCA, pomocí něhož může uživatel vyjádřit prioritu atributů a na jejím základě následně omezit velikost příslušného konceptuálního svazu. AD formulím je věnována kapitola 3.

Jiným problémem, vedle potenciálně obrovského množství formálních konceptů, se kterým se uživatel FCA může setkat, jsou neúplná data. K neúplnosti dat může docházet z různých důvodů, data mohla například vymizet stářím, mohlo dojít k chybě při jejich přenosu, mohla být fyzicky poškozena nebo při jejich sestavování jednoduše všechny informace nebyly známy. Chybí-li v tabulce hodnota 0/1, můžeme si na jejím místě představit nějakou třetí hodnotu, například otazník, kterou interpretujeme jako chybějící informaci. Protože FCA v základní podobě dokáže zpracovávat pouze binární data, bylo by před jejím použitím potřeba odstranit každý řádek nebo sloupec obsahující byť jen jediný otazník, což by znamenalo zbytečnou ztrátu známých informací.

To motivovalo vznik takzvané podmíněné množiny [2]. Za pomoci podmíněných množin lze rozšířit definice různých metematických pojmů tak, aby umožňovaly práci s neúplnými daty. Na základě podmíněných množin vznikla i varianta FCA pro analýzu „dat s otazníky“, tuto variantu nazývám podmíněná FCA. O podmíněných množinách pojednává kapitola 4, podmíněné FCA je věnována podkapitola 4.5.

Cílem této práce byla implementace AD formulí v podmíněné FCA. Výsledky, ke kterým jsem došel, jsou shrnuty v kapitole 5. V podkapitole 5.1 je definována vlastnost podmíněných konceptuálních svazů (ekvivalent konceptuálních svazů v podmíněné FCA), kterou nazývám stabilita. Na rozdíl od klasické FCA, kde existuje jediný konceptuální svaz, podmíněná FCA dává uživateli možnost volby podmíněného konceptuálního svazu. Stabilita se ukázala jako klíčová vlastnost, kterou by měl mít každý podmíněný konceptuální svaz, má-li uživatel v plánu omezovat jej AD formulemi. Dále je v této podkapitole prezentován algoritmus, který daný podmíněný konceptuální svaz učiní stabilním.

Stabilita podmíněných konceptuálních svazů dále umožňuje snadno zobecnit a využívat některé důležité výsledky z teorie klasických AD formulí, což je demonstrováno v podkapitole 5.2.

2 Formální konceptuální analýza

V této kapitole jsou shrnuty pojmy z formální konceptuální analýzy (FCA) potřebné k porozumění následujících kapitol, obsah byl převzat z učebního textu [1].

Jak bylo řečeno v úvodu, FCA je metoda analýzy dat ve tvaru binární relace mezi dvěma množinami. Taková data nazýváme formální kontext.

Definice 1 (formální kontext)
Formální kontext je trojice (X, Y, I), kde X a Y jsou neprázdné množiny a I je binární relace mezi X a Y.

Uvažujme formální kontext (X, Y, I). Prvky množiny X nazýváme objekty, prvky množiny Y nazýváme atributy. Pokud pro $x \in X$, $y \in Y$ platí $(x, y) \in I$, říkáme, že objekt x má atribut y, v opačném případě říkáme, že jej nemá.

Definice 2 (indukované operátory $\uparrow$, $\downarrow$)
Každý formální kontext (X, Y, I) indukuje operátory typu $^{\uparrow} : 2^X \to 2^Y$ a $^{\downarrow} : 2^Y \to 2^X$, které jsou definované pro všechna $A \subseteq X$, $B \subseteq Y$ následovně:

$$A^{\uparrow} = \{y \in Y \mid \text{pro všechna } x \in A \text{ platí } (x, y) \in I\},$$
$$B^{\downarrow} = \{x \in X \mid \text{pro všechna } y \in B \text{ platí } (x, y) \in I\}.$$

Definice 3 (formalní koncept)
Formální koncept v (X, Y, I) je dvojice (A, B) množin $A \subseteq X$ a $B \subseteq Y$, pro které platí $A^{\uparrow} = B$ a $B^{\downarrow} = A$. Množiny A a B nazýváme pořadě extent a intent formálního konceptu (A, B).

Formální koncepty představují pojmy obsažené v datech. Uvažujme například formální kontext, ve kterém objekty jsou živočichové a atributy tvoří vlastnosti typu „má srst“, „živí se masem“, „dýchá žábrami“ a tak podobně. V takovém kontextu přirozeně vznikne například formální koncept odpovídající pojmu „savec“; jeho extentem bude množina všech savců a intentem bude množina všech jim společných vlastností.

Množinu všech formálních konceptů v (X, Y, I), tedy množinu $\{(A, B) \in 2^X \times 2^Y \mid A^{\uparrow} = B \text{ a } B^{\downarrow} = A\}$, běžně označujeme $\mathcal{B}(X, Y, I)$. Formální koncepty můžeme přirozeným způsobem uspořádat.

Definice 4
Definujeme binární relaci $\leq$ na $\mathcal{B}(X, Y, I)$ pro všechny formální koncepty

$(A, B), (C, D) \in \mathcal{B}(X, Y, I)$ následovně:

$$(A, B) \leq (C, D) \text{ právě tehdy, když } A \subseteq C \text{ (právě tehdy, když } B \supseteq D).$$

Je vidět, že relace $\leq$ je částečné uspořádání. Větší formální koncepty odpovídají obecnějším pojmům, zatímco ty menší odpovídají pojmům specifičtějším. Kdyby například v kontextu z předchozí úvahy existovaly formální koncepty C_1, C_2 a C_3 odpovídající pořadě pojmům „pes“, „savec“ a „obratlovec“, platilo by $C_1 \leq C_2 \leq C_3$.

Definice 5 (konceptuální svaz)
Množinu $\mathcal{B}(X, Y, I)$ spolu s relací $\leq$ nazýváme *konceptuální svaz*.

Mohutnost formálního konceptuálního svazu může být vzhledem k velikosti formálního kontextu exponenciální. Dá se například ukázat, že pro formální kontext (X, X, I), kde I je relace nerovnosti, platí $|\mathcal{B}(X, X, I)| = 2^{|X|}$.

Věta 6 (hlavní věta o konceptuálních svazech, první část)
Každý konceptuální svaz je úplný svaz, kde infima a suprema jsou dána následovně:

$$\bigwedge_{i \in I} (A_i, B_i) = \left(\bigcap_{i \in I} A_i, (\bigcap_{i \in I} A_i)^{\uparrow} \right) = \left(\bigcap_{i \in I} A_i, (\bigcup_{i \in I} B_i)^{\downarrow\uparrow} \right),$$

$$\bigvee_{i \in I} (A_i, B_i) = \left((\bigcap_{i \in I} B_i)^{\downarrow}, \bigcap_{i \in I} B_i \right) = \left((\bigcup_{i \in I} A_i)^{\uparrow\downarrow}, \bigcap_{i \in I} B_i \right).$$

Hlavní věta o konceptuálních svazech má ještě druhou část, ta ale pro tento text není potřebná. Celé znění věty včetně důkazu je možné najít v [1] na stranách 15 a 16.

3 AD formule

AD formule (AD je zkratka za „attribute dependency“, česky „atributová závislost“), doplňkový nástroj k FCA, byly představeny v [3]. Pojmy a výsledky potřebné v této práci jsou shrnuty zde, obsah byl převzat z [3] a [4].

Řekněme, že uživatel FCA vnímá některé atributy jako důležitější než jiné. Pro takového uživatele mohou být formální koncepty, jejichž intenty obsahují pouze tyto „méně důležité“ atributy, nezajímavé a (potenciálně velmi početný) konceptuální svaz by se dal pro jeho potřeby zpřehlednit odstraněním těchto „nezajímavých“ formálních konceptů. V takové situaci nacházejí využití AD formule.

Definice 7 (AD formule)

Ať (X, Y, I) je formální kontext. *AD formule* nad množinou atributů Y je výraz tvaru:

$$A \sqsubseteq B,$$

kde A a B jsou podmnožiny množiny Y.

AD formule $A \sqsubseteq B$ vyjadřuje, že „atributy z množiny A jsou méně důležité než atributy z množiny B" nebo přesněji: „formální koncepty, jejichž intenty obsahují nějaký atribut z množiny A, jsou zajímavé jedině tehdy, když obsahují i nějaký atribut z množiny B".

Uvažujme například znovu formální kontext s živočichy a jejich vlastnostmi. Jedna z vlastností by mohla znít takto: „má bílou barvu". Pak by samozřejmě vznikl formální koncept C odpovídající pojmu „bílý živočich". V intentu tohoto formálního konceptu by pravděpodobně ležela pouze tato jediná vlastnost „má bílou barvu", zatímco jeho extent by obsahoval celou řadu rozmanitých živočichů včetně ledního medvěda a různých druhů motýlů.

Uživatel, který data analyzuje například za účelem sestavení taxonomie, by formální koncept C mohl vnímat jako neužitečný, protože kromě bílé barvy jsou živočichové z extentu C naprosto rozdílní. Na druhou stranu formální koncept C' odpovídající pojmu „bílá kočkovitá šelma", jehož intent kromě vlastnosti „má bílou barvu" obsahuje i vlastnosti společné všem kočkovitým šelmám, už by mohl vnímat jako užitečný.

Definice 8

Řekneme, že formální koncept (C, D) v (X, Y, I) splňuje AD formuli $A \sqsubseteq B$, značeno $(C, D) \models A \sqsubseteq B$, jestliže platí následující podmínka:

$$\text{pokud } D \cap A \neq \emptyset, \text{ pak i } D \cap B \neq \emptyset.$$

Požadovaná priorita atributů z předchozí úvahy by se dala vyjádřit AD formulí:

$$\phi = \{\text{„má bílou barvu"}\} \sqsubseteq \{v \mid v \text{ je vlastnost všech kočkovitých šelem}\},$$

pak by platilo $C \not\models \phi$ a $C' \models \phi$. Ovšem bílé kočkovité šelmy nemusí bý jedinou zajímavou skupinou bílých živočichů, stejně tak by uživatel mohl chtít zachovat například formální koncept odpovídající pojmu „bílé vodní ptactvo".

Za tímto účelem není potřeba rozšiřovat množinu atributů na pravé straně ϕ, místo toho je možné navrhnout nové AD formule a formální koncepty filtrovat na základě všech z nich. Tak se zachová přehlednost a systematičnost vyjadřovaných atributových závislostí.

Definice 9

Pro množinu T AD formulí nad Y řekneme, že formální koncept (C, D) splňuje T, značeno $(C, D) \models T$, jestliže splňuje všechny AD formule z T.

Definice 10 (omezený konceptuální svaz)
Množinu všech formálních konceptů v (X, Y, I), které splňují T, označujeme $\mathcal{B}_T(X, Y, I)$ a nazýváme ji *konceptuální svaz omezený množinou* T.

Na rozdíl od klasických formálních konceptuálních svazů, množina $\mathcal{B}_T(X, Y, I)$, přestože ji nazýváme svazem, obecně svazem (ve smyslu svazově uspořádané množiny) být nemusí. Následují výsledky o vlastnostech $\mathcal{B}_T(X, Y, I)$, které opět uvádím bez důkazů, ty je možné najít v [4] na stranách 4 a 5.

Věta 11
$\mathcal{B}_T(X, Y, I)$ je vždy zdola omezený.

Věta 12
Pokud neexistuje $A \sqsubseteq B \in T$ taková, pro kterou by platilo, že A obsahuje nějaký atribut sdílený všemi objekty a zároveň v B žádný takový atribut neleží, pak je $\mathcal{B}_T(X, Y, I)$ i shora omezený.

Věta 13
Pokud jsou všechny AD formule z T tvaru $\{y\} \sqsubseteq \{y'\}$ pro nějaká $y, y' \in Y$, pak je $\mathcal{B}_T(X, Y, I)$ úplný svaz, který je navíc $\bigvee$-podsvazem množiny $\mathcal{B}(X, Y, I)$.

Následující věta říká, že z formálního konceptuálního svazu $\mathcal{B}(X, Y, I)$ mohou být pomocí vhodně sestavené množiny AD formulí T extrahovány formální koncepty tak, že uspořádaná množina $(\mathcal{B}_T(X, Y, I), \leq)$ (uspořádaná relací $\leq$ pro řazení formálních konceptů), odebereme-li nejmenší formální koncept, bude stromem.

Věta 14
Pokud pro každou dvojici atributů $y, y' \in Y$, pro které platí: $\{y\}^{\downarrow} \cap \{y'\}^{\downarrow} \neq \emptyset$ (tedy pro každou dvojici nedisjunktních atributů), obsahuje množina T AD formuli $A \sqsubseteq B$, která splňuje buď:

$$\begin{aligned} &y \in A,\ y' \in B \text{ a všechny atributy z } B \text{ jsou disjunktní, nebo} \\ &y' \in A,\ y \in B \text{ a všechny atributy z } B \text{ jsou disjunktní,} \end{aligned}$$

pak uspořádaná množina $(\mathcal{B}_T(X, Y, I) \setminus \{(Y^{\downarrow}, Y)\}, \leq)$ je strom.

4 Podmíněná množina

V článku [2] byl v rámci obecně vybudované teorie uveden matematický objekt nazývaný podmíněná množina. Tato kapitola obsahuje vybrané definice a výsledky z tohoto článku.

Pro čtenáře obeznámeného s konceptem fuzzy množiny bude obsah této kapitoly lehce uchopitelný. Každá podmíněná množina je totiž technicky vzato fuzzy množinou, rozdíl mezi těmito pojmy spočívá v tom, jakým způsobem je interpretujeme. Uvažujme neprázdné univerzum X. Podmíněná množina A v X je zobrazení typu $X \to L$, kde L je (úplná atomická) Booleova algebra, takzvaná Booleova algebra podmínek, zkráceně BA podmínek. Prvky BA podmínek, které nazýváme podmínky, určitým způsobem modelují neznámou informaci; každá podmínka (kromě prvků $0, 1$) reprezentuje část neznámé informace binárního typu (ta informace může být buď pravdivá nebo nepravdivá). Pokud pro $x \in X$ platí $A(x) = 1$ ($A(x) = 0$), pak víme, že $x \in A$ ($x \notin A$); pokud ovšem platí $0 < A(x) < 1$, pak je informace o příslušnosti prvku x do množiny A neznámá a prvek $A(x)$ interpretujeme jako podmínku, za které $x \in A$.

4.1 Vybrané pojmy z teorie Booleových algeber

K porozumění následující podkapitoly je potřebná základní znalost Booleových algeber. Důkladně vybudovanou teorii na toto téma lze najít například v [5]. Vybrané pojmy definuji zde, definice byly převzaty z [5] a [2].

Definice 15 (Booleova algebra)
Booleova algebra, zkráceně BA, je neprázdná množina G spolu se dvěma význačnými prvky $0, 1 \in G$, dvěma binárními operacemi $\wedge, \vee$ na G a jednou unární operací $\neg$ na G; dále platí, že $\wedge$ a $\vee$ jsou komutativní, vzájemně distributivní a splňují:

- $a \wedge \neg a = 0, \quad a \vee \neg a = 1$ (komplementaritu),
- $a \wedge 1 = a \vee 0 = a$ (neutralitu vůči 1, respektive 0).

Operace $\wedge$ a $\vee$ nazýváme pořadě *průsek* a *spojení*, operaci $\neg$ nazýváme *komplement*.

Definice 16 (indukované uspořádání a operace)
Každá BA G indukuje binární relaci $\leq$ na G:

$$a \leq b \text{ právě tehdy, když } a \wedge b = a \text{ (právě tehdy, když } a \vee b = b)$$

a dvě binární operace $\to$ a $\leftrightarrow$ na G nazývané pořadě *residuum* a *biresiduum*:

$$a \to b = \neg a \vee b, \quad a \leftrightarrow b = (a \to b) \wedge (b \to a).$$

Definice 17 (residuovaný svaz)
Residuovaný svaz je svazově uspořádaná množina L, její nejmenší a největší prvek označujeme pořadě $0, 1 \in L$; na L jsou dále definované dvě binární operace $\otimes, \to$, které splňují:

- $\otimes$ je komutativní, asociativní a neutrální vůči 1,
- operace $\otimes, \to$ jsou provázány takzvanou *podmínkou adjunkce*:

$$a \otimes b \leq c \text{ právě tehdy, když } a \leq b \to c.$$

Operace $\otimes$ a $\to$ nazýváme pořadě *multiplikace* a *residuum*.

Každá BA je residuovaný svaz [2]. Požadovanému uspořádání odpovídá indukované uspořádání a operaci residua odpovídá indukovaná operace residua, multiplikaci pak odpovídá operace průseku.[1]

Definice 18 (homomorfismus, isomorfismus)
Zobrazení $f : G \to H$ mezi BA G, H nazveme *homomorfismus*, jestliže splňuje:

- $f(a \wedge b) = f(a) \wedge f(b)$,
- $f(a \vee b) = f(a) \vee f(b)$,
- $f(\neg a) = \neg f(a)$.

Homomorfismus nazveme *úplný*, jestliže pro všechna $A \subseteq G$ platí:

$$h(\bigvee A) = \bigvee h(A), \quad h(\bigwedge A) = \bigwedge h(A).$$

Bijektivní homomorfismus nazýváme *isomorfismus*. Booleovy algebry nazveme *isomorfní*, jestliže mezi nimi existuje nějaký isomorfismus.

Definice 19 (úplnost a atomicita)
BA G nazveme *úplnou*, je-li $\leq$ úplné svazové uspořádání; dále G nazveme *atomickou*, jestliže pro každé nenulové $b \in G$ existuje atom $a \in G$ (atom je prvek G, který je minimální a přitom nenulový), pro který platí $a \leq b$.

4.2 Booleova algebra podmínek

Nyní zpátky k teorii podmíněnych množin. Jak již bylo řečeno, oborem hodnot každé podmíněné množiny je BA, nazýváme ji Booleova algebra podmínek.

Definice 20 (Booleova algebra podmínek)
Booleova algebra podmínek je BA, která je úplná a atomická.

Přestože podmíněné množiny (a tedy i BA podmínek) jsou obecný nástroj s možným využitím i v jiných oblastech [2], v tomto textu se soustředím na využití v FCA. Představme si formální kontext „s otazníky“, jako byl popsán v úvodu (kromě nul a jedniček se v tabulce mohou vyskytovat otazníky reprezentující neznámou informaci 0/1). Označme n počet otazníků. Každý otazník představuje buď hodnotu 0 nebo hodnotu 1, existuje tedy 2^n úplných formálních kontextů (bez otazníků), z nichž pouze jediný odpovídá skutečnosti.

[1] Residuované svazy, ve kterých multiplikace takto koinciduje s průsekem, nazýváme Heytingovy algebry.

Mohlo by se ovšem stát, že mezi neznámými hodnotami na pozicích otazníků existují nějaké závislosti. Například: „pokud objekt x má atribut y, pak objekt x' atribut y jistě nemá“ nebo „objekt x má atribut y_1 právě tehdy, když má alespoň jeden z atributů y_2 nebo y_3“ a tak podobně.

K modelování formálního kontextu „s otazníky“ lze využít podmíněnou množinu $I : X \times Y \to L$, kde X je množina objektů, Y je množina atributů a L je nějaká vhodně sestavená BA podmínek. Prvky $0, 1 \in L$ pak vyjadřují úplnou informaci (nulám a jedničkám v původní tabulce by tedy odpovídaly prvky $0, 1 \in L$). Otazníky nahradí ostatní prvky z L. Výhoda tohoto přístupu spočívá v tom, že případné závislosti mezi neznámými hodnotami je možné vyjádřit s využitím operací komplementu, průseku a spojení.

Příklad 21

Máme následující formální kontext s neúplnou informací:

	y
x_1	?
x_2	?
x_3	?
x_4	?

a víme, že mezi neznámými hodnotami platí následující závislosti:

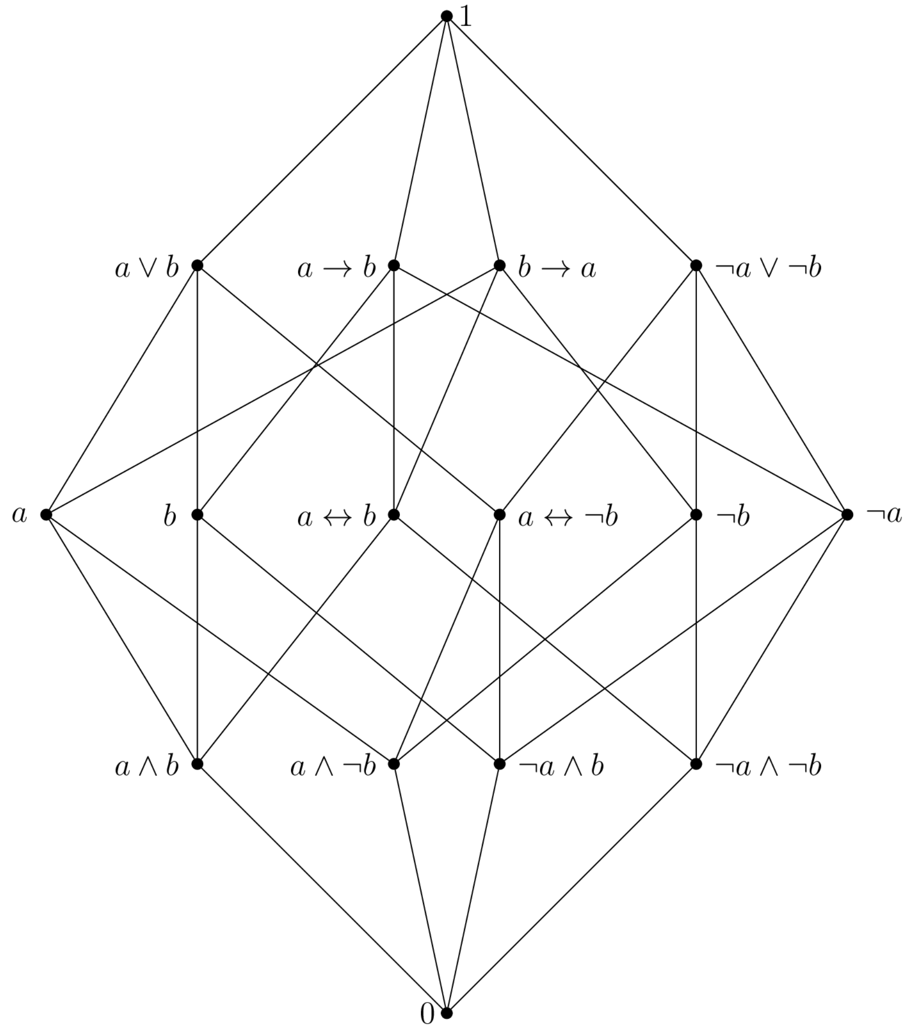
„x_1 má y právě tehdy, když x_2 i x_3 mají y“,
„x_4 má y právě tehdy, když x_1 nemá y“.

Sestavíme BA podmínek L, která bude mimo jiné obsahovat prvky a, b; tyto prvky budou reprezentovat neznámé informace pořadě „x_2 má y“ a „x_3 má y“. Kromě toho by pro tento příklad stačilo přidat do L prvky 0, 1 a prvky $(a \wedge b)$ a $\neg(a \wedge b) = \neg a \vee \neg b$, kterými bychom mohli reprezentovat neznámé informace „x_1 má y“ a „x_4 má y“ a zachytit tak dané závislosti. My ovšem do L přidáme ještě další relevantní prvky jako $\neg a$ nebo $a \leftrightarrow b$. Celkem bude L obsahovat 16 prvků a je znázorněna na obrázku 1.

Jiný pohled na L: představme si (nekonečnou) množinu F všech formulí výrokové logiky s množinou výrokových proměnných $\{a, b\}$ a množinou logických spojek $\{\neg, \wedge, \vee\}$. Označme F' rozklad F podle sémantické ekvivalence. Pak množina F' společně se třemi operacemi na F', které provádí logickou negaci, konjunkci a disjunkci na libovolně zvolených reprezentantech z daných tříd rozkladu a vrací třídu rozkladu příslušnou výsledné formuli, tvoří BA. Prvku 0 odpovídá třída kontradikcí a prvku 1 odpovídá třída tautologií. Tato BA bude isomorfní s naší BA podmínek L.

Formální kontext namodelovaný s pomocí L potom vypadá následovně:

	y
x_1	$(a \wedge b)$
x_2	a
x_3	b
x_4	$\neg a \vee \neg b$



Obrázek 1: BA podmínek se dvěma elementárními podmínkami a a b.

Uvedu ještě jeden příklad, ve kterém opět figurují závislosti neznámých hodnot. Tentokrát ovšem nejde o závislosti mezi dvěma hodnotami ve formálním kontextu, ale o závislost mezi hodnotou ve formálním kontextu a externí informací mimo formální kontext. Příklad se týká dědičnosti krevních skupin.

Příklad 22 (krevní skupiny)

V následujícím formálním kontextu objekty tvoří lidé Anna, Bert a Cyril, atributy jsou krevní skupiny. Víme, že Cyril je potomkem Anny a Berta. Anna má krevní skupinu A, Bert má krevní skupinu 0. Cyril svou krevní skupinu nezná.

	A	B	AB	0
Anna	1	0	0	0
Bert	0	0	0	1
Cyril	?	?	?	?

Anna má krevní skupinu A, nevíme ovšem, zda od svých rodičů zdědila dvě dominantní alely A, nebo jednu dominantní alelu A a jednu recesivní alelu 0.

Což je skutečnost, která má vliv na krevní skupinu Cyrila. Do BA podmínek pro tento příklad tedy přidáme prvek $Anna\text{-}AA$ reprezentující informaci: „Anna nese alely AA“; prvek $\neg Anna\text{-}AA$ pak můžeme přirozeně reprezentovat jako informaci: „Anna nese alely A0“, protože jiná možnost nastat nemůže.

Další neznámá skutečnost je tato: zdědil Cyril od Anny levou nebo pravou alelu? (Levou a pravou rozumíme vzhledem k zápisu; tedy nese-li Anna alely A0, levou myslíme alelu A a pravou alelu 0.) Přidáme tedy do BA podmínek prvek *left*, kterým reprezentujeme informaci: „Cyril od Anny zdědil levou alelu“.

Z pravidel pro dědění krevních skupin vyvodíme, že Cyril může mít jedině krevní skupinu A nebo 0:

	A	B	AB	0
Anna	1	0	0	0
Bert	0	0	0	1
Cyril	?	0	0	?

S pomocí BA podmínek ovšem můžeme situaci vyjádřit přesněji (pro přehlednost byly vynechány nulové sloupce odpovídající krevním skupinám B a AB):

	A	0
Anna	1	0
Bert	0	1
Cyril	$(Anna\text{-}AA \vee \mathit{left})$	$(\neg Anna\text{-}AA \wedge \neg \mathit{left})$

S případným doplněním informace, byť třeba jen částečným, se v této teorii pracuje přechodem k jiné (menší) BA podmínek. Za tímto účelem se definují takzvané reality.

Definice 23 (realita)

Ať L, K jsou BA podmínek. Surjektivní úplný homomorfismus $h : L \to K$ nazýváme *realita*. Obsahuje-li K pouze dva prvky $0, 1 \in K$, nazveme h *totální realitou*.

Pro realitu $h : L \to K$ a prvek $b \in L$ interpretujeme skutečnost $h(b) = 1$ ($h(b) = 0$) tak, že neznámá informace, kterou modeluje b, je pravdivá (nepravdivá) v situaci, kterou modeluje realita h.

Totální reality tedy modelují doplnění všech neznámých informací. Booleovy algebry podmínek, které obsahují pouze prvky $0, 1$, budeme označovat symbolem 2. Každé totální realitě jednoznačně odpovídá nějaký atom. Přesněji řečeno: pro každou totální realitu $h : L \to 2$ existuje právě jeden atom $a \in L$ tak, že $h(a) = 1$ a naopak pro každý atom z L existuje právě jedna taková totální realita [2].

Podívejme se znovu na BA podmínek na obrázku 1. Jak bylo popsáno v příkladu 21, na její prvky můžeme nahlížet jako na třídy formulí výrokové logiky se dvěma výrokovými proměnnými a, b (popisky u jednotlivých uzlů odpovídají reprezentantům příslušných tříd). Při takové interpretaci korespondují totální reality (atomy) s pravdivostními ohodnoceními výrokových proměnných.

4.3 Podmíněná množina, podmíněná relace

Definice 24 (podmíněná množina)
Pro BA podmínek L a neprázdnou množinu X, *L-podmíněná množina* A v X je libovolné zobrazení $A : X \to L$.

Namísto „L-podmíněná množina“ budu často psát jen „podmíněná množina“, nebude-li potřeba explicitně zmiňovat BA podmínek L.

S využitím podmíněných množin lze rozšířit definice různých matematických pojmů tak, aby umožňovaly práci s neúplnou informací. Neúplnou informaci vždy zastupuje nějaká BA podmínek L. Takto rozšířené pojmy pak budu označovat jako „L-podmíněné“, přičemž předponu „L-“ budu, pro lepší čitelnost, mnohdy vynechávat. Původní (nerozšířené) pojmy budu označovat jako „nepodmíněné“.

Každý podmíněný objekt většinou zastupuje hned několik nepodmíněných objektů, přičemž každý ze zastupovaných nepodmíněných objektů odpovídá jinému zúplnění informace. Například podmíněná množina A v X, pro kterou existuje jediné $x_0 \in X$ takové, že $0 < A(x_0) < 1$, zastupuje dvě nepodmíněné množiny; v obou z nich leží právě ta $x \in X$, pro která platí $A(x) = 1$ a v jedné leží navíc i prvek x_0.

Každý podmíněný objekt můžeme s každým doplněním informace, tedy s každou realitou, zpřesnit (zmenšit počet jím zastupovaných nepodmíněných objektů), takové zpřesňování nazýváme realizace. Pro podmíněný objekt X a realitu h, realizace X podle h je (přesnější) podmíněný objekt označovaný X^h. Následuje definice realizace podmíněné množiny.[2]

Pro všechny následující definice až do konce kapitoly uvažujme Booleovy algebry podmínek L, K.

Definice 25 (realizace podmíněné množiny)
Realizace L-podmíněné množiny A v X podle reality $h : L \to K$ je K-podmíněná množina A^h v X, která pro všechna $x \in X$ splňuje:

$$A^h(x) = h(A(x)).$$

Rozšířením pojmu relace je takzvaná podmíněná relace. S podmíněnou relací jsme se už setkali v předchozí podkapitole, když jsme modelovali formální kontext „s otazníky“, vztahy mezi prvky množin X a Y, pomocí podmíněné množiny v $X \times Y$.

Definice 26 (podmíněná relace)
Pro neprázdné množiny $X_1, \ldots, X_n$, *L-podmíněná relace* mezi $X_1, \ldots, X_n$ je libovolná L-podmíněná množina v $X_1 \times \ldots \times X_n$.

[2] Tuto definici v následující podkapitole ještě zobecním. Jde o to, že vhodně rozšířit lze i definiční obor podmíněných množin. K tomuto rozšíření je ovšem potřeba nejprve definovat podmíněnou rovnost, kterou definuji na konci podkapitoly.

Podobně jako u klasických relací, pokud je v předchozí definici $n = 2$, mluvíme o binární podmíněné relaci; pokud navíc $X_1 = X_2$, tedy pokud jde o podmíněnou množinu v X^2, pak mluvíme o binární podmíněné relaci na X.

Definice 27 (podmíněná rovnost)
Binární L-podmíněnou relaci $\approx$ na množině X nazveme *L-podmíněná rovnost*, jestliže splňuje:

- $x \approx x = 1$ (reflexivitu),
- $x_1 \approx x_2 = x_2 \approx x_1$ (symetrii),
- $(x_1 \approx x_2) \wedge (x_2 \approx x_3) \leq x_1 \approx x_3$ (tranzitivitu),
- $x_1 \approx x_2 = 1$ implikuje $x_1 = x_2$ (separabilitu).

4.4 Podmíněné univerzum

Podmíněné množiny umožňují, prostřednictvím prvků BA podmínek, vyjádřit neznámou informaci ohledně příslušnosti prvku univerza do množiny. Jiným typem neznámé informace je potom neznalost toho, zda se dva prvky univerza rovnají, nebo jsou odlišné. Aby bylo možné s podmíněnými množinami pracovat i nad daty s neúplnou informací ohledně rovnosti prvků, zavádí se takzvaná podmíněná univerza.

Definice 28 (podmíněné univerzum)
Neprázdnou množinu X společně s L-podmíněnou rovnosí $\approx$ na X nazveme *L-podmíněné univerzum*.

Definice 29 (isomorfismus podmíněných univerz)
Dvě L-podmíněná univerza $(X, \approx_X)$, $(Y, \approx_Y)$ nazveme *isomorfní*, jestliže existuje bijektivní zobrazení $f : X \to Y$, které splňuje:

$$f(x) \approx_Y f(x') = x \approx_X x'.$$

Definice 30 (realizace podmíněného univerza)
Realizace L-podmíněného univerza $(X, \approx_X)$ podle reality $h : L \to K$ je surjektivní zobrazení $f : X \to Y$, kde Y je nosná množina nějakého K-podmíněného univerza $(Y, \approx_Y)$, přičemž platí:

$$f(x) \approx_Y f(x') = h(x \approx_X x').$$

Dá se ukázat, že realizace podmíněného univerza je jednoznačně určena realitou. Přesněji řečeno: pro dvě realizace $f_1 : X \to Y_1$, $f_2 : X \to Y_2$ L-podmíněného univerza $(X, \approx_X)$ podle reality $h : L \to K$ platí, že příslušná K-podmíněná univerza $(Y_1, \approx_{Y_1})$, $(Y_2, \approx_{Y_2})$ jsou isomorfní [2].

Nyní uvažujme L-podmíněné univerzum $(X, \approx_X)$ a realitu $h : L \to K$. Označme $\approx^1$ relaci ekvivalence na X danou takto: $x \approx^1 x'$ právě tehdy, když $h(x \approx_X x') = 1$. Dále označme $X\backslash_{\approx^1}$ rozklad množiny X podle $\approx^1$. Pak zobrazení $f : X \to X\backslash_{\approx^1}$, které každému prvku $x \in X$ přiřazuje jeho třídu rozkladu $[x]_{\approx^1} \in X\backslash_{\approx^1}$, je realizací $(X, \approx_X)$ podle h. Příslušné K-podmíněné univerzum je pak tvaru $(X\backslash_{\approx^1}, \approx_X^h)$, kde K-podmíněná rovnost $\approx_X^h$ pro všechny třídy rozkladu $[x]_{\approx^1}, [x']_{\approx^1} \in X\backslash_{\approx^1}$ splňuje $[x]_{\approx^1} \approx_X^h [x']_{\approx^1} = h(x \approx_X x')$ [2]. Takto definované zobrazení vždy existuje pro libovolné L-podmíněné univerzum a libovolnou realitu $h : L \to K$.

Na základě předchozích úvah, kdykoliv bude v tomto textu řeč o realizaci L-podmíněného univerza $(X, \approx_X)$ podle reality $h : L \to K$, budeme předpokládat, že máme k dispozici nějakou realizaci $f : X \to Y$ L-podmíněného univerza $(X, \approx_X)$ podle h, přičemž samotné zobrazení f často nebudeme blíže specifikovat; není to potřeba, protože všechna z nich by ve výsledku určovala (až na isomorfismus) stejná K-podmíněná univerza. Realizací $(X, \approx_X)$ podle $h : L \to K$ pak budeme nazývat samotné K-podmíněné univerzum $(Y, \approx_Y)$, které budeme značit takto $(X^h, \approx_X^h)$. Obraz $f(x)$ prvku $x \in X$ budeme značit $(x)^h$ nebo jen x^h.

Definice 31 (součin podmíněných univerz)

Pro L-podmíněná univerza $(X, \approx_X)$, $(Y, \approx_Y)$ definujeme *součin L-podmíněných univerz* $(X, \approx_X) \times (Y, \approx_Y)$ jako množinu $X \times Y$ spolu s L-podmíněnou rovností $\approx_{X \times Y}$, která je dána takto:

$$(x, y) \approx_{X \times Y} (x', y') = (x \approx_X x') \wedge (y \approx_Y y').$$

Součin L-podmíněných univerz je L-podmíněné univerzum [2]. V předchozí podkapitole jsme podmíněné množiny definovali jako zobrazení z klasické množiny do nějaké BA podmínek. U každé klasické množiny ovšem vždy předpokládáme úplnou znalost ohledně rovnosti prvků. Dalo by se říct, že ke každé klasické množině X vždy implicitně uvažujeme i relaci identity id na X. Na dvojici (X, id) můžeme nahlížet jako na speciální typ podmíněného univerza, ve kterém podmíněná rovnost koinciduje s identitou. Odteď budeme u L-podmíněných množin jakožto definiční obory připouštět obecná L-podmíněná univerza. Následuje avizované zobecnění definice realizace podmíněné množiny.

Definice 32 (realizace podmíněné množiny)

Realizace L-podmíněné množiny A v $(X, \approx)$ podle reality $h : L \to K$ je K-podmíněná množina A^h v $(X^h, \approx^h)$, která pro všechna $x^h \in X^h$ splňuje:

$$A^h(x^h) = \bigvee_{\substack{x' \in X: \\ (x')^h = x^h}} h(A(x')).$$

Důležitým pojmem, který se týká podmíněných množin v podmíněných univerzech, je extensionalita, nebo-li kompatibilita s podmíněnou rovností.

Definice 33 (extensionalita)
L-podmíněnou množinu A v $(X, \approx)$ nazveme *extensionální*, nebo taky *kompatibilní s* $\approx$, jestliže splňuje:

$$A(x) \wedge (x \approx x') \leq A(x'). \tag{1}$$

L-podmíněnou relaci R mezi L-podmíněnými univerzy $(X_1, \approx_{X_1}), \ldots, (X_n, \approx_{X_n})$, která je extensionální L-podmíněnou množinou v $(X_1, \approx_{X_1}) \times \ldots \times (X_n, \approx_{X_n})$, tedy která splňuje:

$$R(x_1, \ldots, x_n) \wedge (x_1 \approx_{X_1} x_1') \wedge \ldots \wedge (x_n \approx_{X_n} x_n') \leq R(x_1', \ldots, x_n'),$$

také nazveme *extensionální*, či *kompatibilní s* $\approx_{X_1}, \ldots, \approx_{X_n}$.

Podmínku (1) můžeme přepsat následovně:

$$\begin{aligned} A(x) \wedge (x \approx x') &\leq A(x'), \\ (x \approx x') &\leq A(x) \to A(x') \quad \text{(adjunkce)}, \\ (x \approx x') &\leq A(x') \to A(x) \quad \text{(symetrie } \approx\text{)}, \\ (x \approx x') &\leq (A(x) \to A(x')) \wedge (A(x') \to A(x)), \\ (x \approx x') &\leq A(x) \leftrightarrow A(x'), \end{aligned}$$

přičemž poslední výraz $(x \approx x') \leq A(x) \leftrightarrow A(x')$ můžeme číst takto: pokud jsou si prvky x a x' podobné, pak jsou si podobné i podmínky, za kterých tyto prvky patří do podmíněné množiny A.

Pro libovolnou podmíněnou množinu A v $(X, \approx)$ a libovolnou realitu h platí, že je-li A extensionální, pak její realizace A^h podle h splňuje [2]:

$$A^h(x^h) = h(A(x))$$

4.5 Podmíněná FCA

V předchozích podkapitolách byla řeč o formálním kontextu „s otazníky“ a na příkladech byl demonstrován způsob, jakým lze otazníky s výhodou nahradit prvky vhodně zvolené BA. K výjádření neznámé informace o vztahu mezi objekty a atributy (a případnému zaznamenání závislostí mezi takovými vztahy) se využívala podmíněná relace.

Dále byl uveden jiný typ neznámé informace, totiž neznalost ohledně rovnosti prvků univerza, a bylo představeno podmíněné univerzum jakožto nástroj pro práci nad daty s neúplnou informací tohoto typu. Kromě toho byla uvedena extensionalita, podmínka na podmíněné množiny (nebo podmíněné relace), která zaručuje určitou jejich kompatibilitu s podmíněnými univerzy.

V této podkapitole představíme rozšíření definic základních pojmů z FCA pro práci nad daty s neúplnou informací obou zmíněných typů. Vznikne tak varianta FCA, kterou budu označovat jako podmíněná FCA.

Definice 34 (podmíněný kontext)
L-podmíněný kontext je trojice $((X,\approx_X),(Y,\approx_Y),I)$, kde $(X,\approx_X)$ a $(Y,\approx_Y)$ jsou L-podmíněná univerza a I je extensionální L-podmíněná relace mezi $(X,\approx_X)$ a $(Y,\approx_Y)$.

Nebude-li potřeba zmiňovat podmíněné rovnosti $\approx_X$, $\approx_Y$, budeme podmíněný kontext $((X,\approx_X),(Y,\approx_Y),I)$ zkráceně zapisovat takto (X,Y,I).

Z extensionality libovolné podmíněné množiny A plyne, že i její realizace A^h podle libovolné reality h bude extensionální [2]. Speciálně tedy, zrealizujeme-li všechny složky libovolného podmíněného kontextu $((X,\approx_X),(Y,\approx_Y),I)$ podle reality h, pak výsledná trojice $((X^h,\approx_X^h),(Y^h,\approx_Y^h),I^h)$ bude opět tvořit podmíněný kontext, protože I^h zůstane extensionální.

Definice 35 (realizace podmíněného kontextu)
Realizace L-podmíněného kontextu $((X,\approx_X),(Y,\approx_Y),I)$ podle reality $h : L \to K$ je K-podmíněný kontext $((X^h,\approx_X^h),(Y^h,\approx_Y^h),I^h)$.

Opět, nebude-li potřeba zmiňovat podmíněné rovnosti $\approx_X^h$, $\approx_Y^h$, budeme realizaci $((X^h,\approx_X^h),(Y^h,\approx_Y^h),I^h)$ podmíněného kontextu $((X,\approx_X),(Y,\approx_Y),I)$ podle reality h zkráceně zapisovat takto (X^h,Y^h,I^h).

Stejně jako u každého jiného podmíněného objektu, zrealizujeme-li L-podmíněný kontext (X,Y,I) podle nějaké totální reality $h : L \to 2$, zúplníme tím veškeré informace v (X,Y,I) a na (X^h,Y^h,I^h) pak můžeme nahlížet jako na (nepodmíněný) formální kontext.

Pro následující definice uvažujme L-podmíněný kontext (X,Y,I).

Definice 36 (podmíněný koncept)
Pro extensionální L-podmíněné množiny A v $(X,\approx_X)$, B v $(Y,\approx_Y)$ dvojici (A,B) nazveme *L-podmíněný koncept* v (X,Y,I), jestliže pro každou totální realitu $h : L \to 2$ platí, že (A^h,B^h) je formální koncept v (X^h,Y^h,I^h). Podmíněné množiny A, B pak nazveme pořadě *extent* a *intent* podmíněného konceptu (A,B).

Definice 37 (realizace podmíněného konceptu)
Realizace L-podmíněného konceptu (A,B) v (X,Y,I) podle reality $h : L \to K$ je K-podmíněný koncept (A^h,B^h) v (X^h,Y^h,I^h).

Definice 38 (podmíněné indukované operátory $\uparrow$, $\downarrow$)
Pro libovolné L-podmíněné množiny A v $(X,\approx_X)$, B v $(Y,\approx_Y)$ definujeme L-podmíněné množiny $A^\uparrow$ v $(Y,\approx_Y)$, $B^\downarrow$ v $(X,\approx_X)$ následovně:

$$A^\uparrow(y) = \bigwedge_{x\in X} A(x) \to I(x,y),$$
$$B^\downarrow(x) = \bigwedge_{y\in Y} B(y) \to I(x,y).$$

Pro libovolné extensionální podmíněné množiny A v $(X, \approx_X)$, B v $(Y, \approx_Y)$ platí, že dvojice (A, B) tvoří podmíněný koncept v (X, Y, I) právě tehdy, když $A^{\uparrow} = B$ a $B^{\downarrow} = A$ [2].

Definice 39

Definujeme binární L-podmíněné relace $\preceq$ a $\approx$ na $\mathbb{B}^*(X, Y, I)$ pro všechny L-podmíněné koncepty $(A, B), (C, D) \in \mathbb{B}^*(X, Y, I)$ následovně:

$$(A, B) \preceq (C, D) = \bigwedge_{x \in X} A(x) \to C(x) \left(= \bigwedge_{y \in Y} D(y) \to B(y)\right),$$
$$(A, B) \approx (C, D) = ((A, B) \preceq (C, D)) \wedge ((C, D) \preceq (A, B)).$$

Pro množinu podmíněných konceptů $V \subseteq \mathbb{B}^*(X, Y, I)$ a realitu h budeme označením V^h mínit množinu realizací všech podmíněných konceptů z V podle reality h, tedy:

$$V^h = \{(A^h, B^h) \mid (A, B) \in V\}.$$

Definice 40 (podmíněný konceptuální svaz)

Libovolnou množinu $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ L-podmíněných konceptů v (X, Y, I) spolu s L-podmíněnou relací $\preceq$ nazveme *L-podmíněný konceptuální svaz*, jestliže pro libovolnou totální realitu $h : L \to 2$ platí, že $\mathbb{B}^h$ spolu s $\preceq^h$ tvoří konceptuální svaz v (X^h, Y^h, I^h).

Příklad 41

Označme L BA podmínek z obrázku 1 a uvažujme L-podmíněná univerza $(X = \{x_1, x_2\}, \approx_X)$, $(Y = \{y_1, y_2\}, \approx_Y)$, kde $\approx_Y$ odpovídá obyčejné identitě (tedy platí: $y_1 \approx_Y y_2 = 0$) a pro $\approx_X$ platí: $x_1 \approx_X x_2 = a$.

Dále uvažujme následující L-podmíněný kontext (X, Y, I):

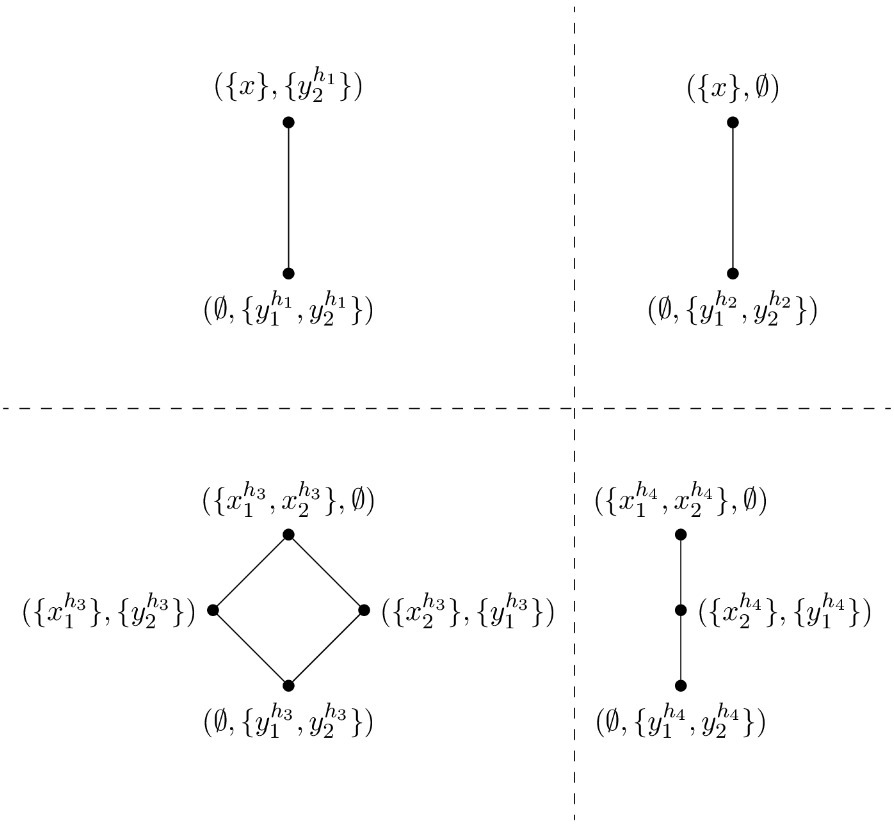
	y_1	y_2
x_1	0	b
x_2	$\neg a$	$a \wedge b$

Existují čtyři totální reality typu $L \to 2$, označme je h_1, h_2, h_3 a h_4. V každé z těchto totálních realit je pravdivý právě jeden atom z L; řekněme, že platí:

$$h_1(a \wedge b) = 1, \quad h_2(a \wedge \neg b) = 1, \quad h_3(\neg a \wedge b) = 1, \quad h_4(\neg a \wedge \neg b) = 1.$$

Na obrázku 2 jsou znázorněny konceptuální svazy $\mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})$ pro každou totální realitu h_i. Na obrázku 3 jsou potom znázorněny dva vybrané podmíněné konceptuální svazy $\mathbb{B}^1, \mathbb{B}^2 \subseteq \mathbb{B}^*(X, Y, I)$. Čtenář si může snadno ověřit, že jde opravdu o podmíněné konceptuální svazy, tedy že pro každou totální realitu h_i platí: $(\mathbb{B}^1)^{h_i} = \mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i}) = (\mathbb{B}^2)^{h_i}$.

Ve zbytku kapitoly uvedu větu, na kterou se budu často odvolávat v následující kapitole. Jedná se o důsledek jednoho z výsledků z [2] (Theorem 11, strana 26), zde se ale omezujeme na konečné BA podmínek.



Obrázek 2: Konceptuální svazy příslušné realizacím podmíněného kontextu (X, Y, I) z příkladu 41 podle všech totálních realit. Prvky $x_1^{h_1} = x_2^{h_1}$ a $x_1^{h_2} = x_2^{h_2}$ jsou označeny x. Nahoře vlevo: $\mathcal{B}(X^{h_1}, Y^{h_1}, I^{h_1})$, nahoře vpravo: $\mathcal{B}(X^{h_2}, Y^{h_2}, I^{h_2})$, dole velvo: $\mathcal{B}(X^{h_3}, Y^{h_3}, I^{h_3})$, dole vpravo: $\mathcal{B}(X^{h_4}, Y^{h_4}, I^{h_4})$.

Uvažujme konečnou BA podmínek L a L-podmíněný kontext (X, Y, I). Protože L je konečná, obsahuje konečný počet atomů, a tedy existuje i konečný počet totálních realit $h_1, \ldots h_n : L \to 2$. Označme $a_i \in L$ atom příslušný totální realitě h_i (tedy $h_i(a_i) = 1$ a $h_i(a_j) = 0$ pro všechna $j \neq i$). Dále označme $\mathcal{B}^1 \times \ldots \times \mathcal{B}^n$ kartézský součin všech konceptuálních svazů $\mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})$ pro $i \in \{1, \ldots, n\}$, tedy:

$$\mathcal{B}^1 \times \ldots \times \mathcal{B}^n = \mathcal{B}(X^{h_1}, Y^{h_1}, I^{h_1}) \times \ldots \times \mathcal{B}(X^{h_n}, Y^{h_n}, I^{h_n}).$$

Formální koncepty z konceptuálního svazu $\mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})$ budeme označovat C_j^i, kde horní index $i \in \{1, \ldots, n\}$ odkazuje k totální realitě h_i a dolní index $j \in \{0, \ldots, m-1\}$ pro $m = |\mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})|$ rozlišuje mezi jednotlivými formálními koncepty v tomto konceptuálním svazu, přičemž C_0^i bude vždy označovat nejmenší formální koncept v $\mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})$. Extent a intent formálního konceptu $C_j^i \in \mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})$ budeme označovat $Ext(C_j^i)$ a $Int(C_j^i)$.

Věta 42
Množina $\mathcal{B}^1 \times \ldots \times \mathcal{B}^n$ je ve vzájemně jednoznačném vztahu s množinou $\mathbb{B}^*(X, Y, I)$ všech podmíněných konceptů v (X, Y, I). Příslušné bijektivní zobra-

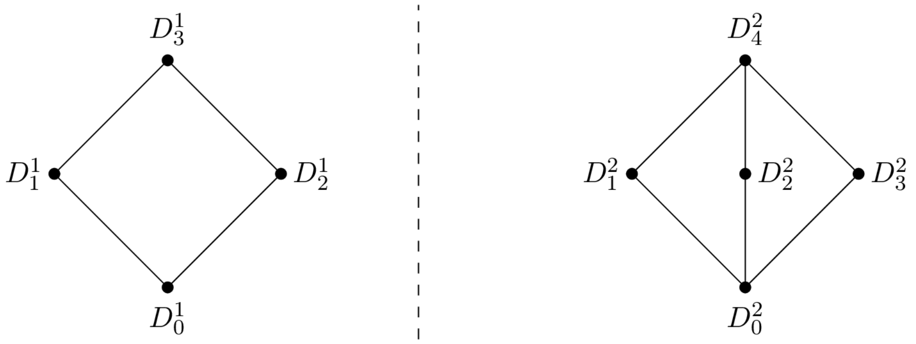
zení $f : \mathcal{B}^1 \times \ldots \times \mathcal{B}^n \to \mathbb{B}^*(X, Y, I)$ přiřazuje n-ticím $(C^1_{x_1}, \ldots, C^n_{x_n})$ formálních konceptů podmíněné koncepty $f(C^1_{x_1}, \ldots, C^n_{x_n}) = (A, B) \in \mathbb{B}^*(X, Y, I)$, pro které platí:

$$A(x) = \bigvee_{x \,\in\, Ext(C^i_{x_i})} a_i,$$

$$B(y) = \bigvee_{y \,\in\, Int(C^i_{x_i})} a_i.$$

Navíc pro každou totální realitu h_i a každou n-tici $(C^1_{x_1}, \ldots, C^n_{x_n})$ formálních konceptů platí:

$$(f(C^1_{x_1}, \ldots, C^n_{x_n}))^{h_i} = C^i_{x_i}.$$



$D^1_0 = (\emptyset, \{y_1, y_2\})$
$D^1_1 = (\{^{\neg a}/x_2\}, \{y_1, ^{a}/y_2\})$
$D^1_2 = (\{^{b}/x_1, ^{a\wedge b}/x_2\}, \{^{\neg b}/y_1, y_2\})$
$D^1_3 = (\{x_1, x_2\}, \{^{a\wedge b}/y_2\})$

$D^2_0 = (\emptyset, \{y_1, y_2\})$
$D^2_1 = (\{^{\neg a}/x_2\}, \{y_1, ^{a}/y_2\})$
$D^2_2 = (\{^{\neg a\vee\neg b}/x_1, ^{\neg a\vee\neg b}/x_2\}, \{^{a\wedge b}/y_1, ^{a\wedge b}/y_2\})$
$D^2_3 = (\{x_1, x_2\}, \{^{a\wedge b}/y_2\})$

Obrázek 3: Podmíněné konceptuální svazy $\mathbb{B}^1$ (vlevo) a $\mathbb{B}^2$ (vpravo) příslušné podmíněnému kontextu (X, Y, I) z příkladu 41. Na obrázcích jsou podmíněné koncepty uspořádány relací $\preceq^1$, pro kterou platí: $C \preceq^1 D$ právě tehdy, když $C \preceq D = 1$. Podmíněné množiny odpovídající extentům a intentům jsou dány výčtem prvků podobně, jako se zapisují fuzzy množiny (podmínka nad lomítkem odpovídá podmínce, za které příslušný objekt nebo atribut patří do množiny).

5 AD formule v podmíněné FCA

V této poslední kapitole jsou shrnuty všechny výsledky, ke kterým jsem došel při snaze o implementaci AD formulí v podmíněné FCA. Omezil jsem se přitom na konečné BA podmínek; některé výsledky nicméně předpoklad konečnosti BA podmínek nevyžadují, konkrétně jsou to ty výsledky, ve kterých se neodvolávám na Větu 42. Až do konce textu tedy uvažujme konečnou BA podmínek L a L-podmíněný kontext (X, Y, I).

Definice 43 (AD formule)

AD formule nad L-podmíněným univerzem $(Y, \approx_Y)$ je výraz tvaru

$$A \sqsubseteq B,$$

kde A a B jsou podmnožiny množiny Y.

AD formulemi v podmíněné FCA tedy rozumíme stejné výrazy jako v klasické FCA, jedniný rozdíl spočívá v tom, že nyní jde o množiny atributů z podmíněného univerza; připouštíme tedy možnost neúplné informace ohledně rovnosti atributů. Proto je potřeba zavést realizaci AD formule.

Definice 44 (realizace AD formule)

Realizace AD formule $\phi = \{y_i, \ldots, y_j\} \sqsubseteq \{y_k, \ldots, y_l\}$ nad $(Y, \approx_Y)$ podle reality $h : L \to K$ je AD formule $\phi^h = \{y_i^h, \ldots, y_j^h\} \sqsubseteq \{y_k^h, \ldots, y_l^h\}$ nad $(Y^h, \approx_Y^h)$.

Definice 45

Řekneme, že L-podmíněný koncept (A, B) v (X, Y, I) splňuje AD formuli ϕ nad $(Y, \approx_Y)$, značeno $(A, B) \models \phi$, jestliže pro každou totální realitu $h : L \to 2$ platí $(A^h, B^h) \models \phi^h$. Dále uvažujme množinu T AD formulí nad $(Y, \approx_Y)$. Řekneme, že (A, B) splňuje T, značno $(A, B) \models T$, jestliže (A, B) splňuje všechny AD formule z T.

Definice 46 (omezený podmíněný konceptuální svaz)

Pro L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ a množinu T AD formulí nad $(Y, \approx_Y)$ nazveme množinu

$$\mathbb{B}_T = \{(A, B) \in \mathbb{B} \mid (A, B) \models T\}$$

L-podmíněný konceptuální svaz $\mathbb{B}$ omezený množinou T.

5.1 Stabilní podmíněné konceptuální svazy

Podmíněná FCA nabízí uživateli možnost volby podmíněného konceptuálního svazu. Pro různé potřeby uživatele mohou být vhodné různé podmíněné konceptuální svazy, které se odlišují tím, jak dobře zachovávají vlastnosti příslušných konceptuálních svazů v jednotlivých totálních realitách. Například v článku [2] se definuje takzvaná supremum kompatibilita: L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ nazveme supremum kompatibilní, jestliže pro každé $W \subseteq \mathbb{B}$ a pro každou totální realitu $h : L \to 2$ platí $(\bigvee W)^h = \bigvee W^h$.

V této podkapitole definuju vlastnost podmíněných konceptuálních svazů, kterou pro množinu T AD formulí nazývám T-stabilita. Tato vlastnost se ukázala být klíčovou při používání AD formulí v podmíněné FCA.

Pro množinu T AD formulí nad L-podmíněným univerzem a realitu $h : L \to K$ budeme označením T^h mínit množinu realizací všech AD formulí z T podle reality h, tedy:

$$T^h = \{\phi^h \mid \phi \in T\}.$$

Věta 47

Pro každý L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X,Y,I)$, každou množinu T AD formulí nad $(Y, \approx_Y)$ a každou totální realitu $h : L \to 2$ platí:

$$(\mathbb{B}_T)^h \subseteq \mathcal{B}_{T^h}(X^h, Y^h, I^h). \tag{2}$$

Důkaz

Pro $D^h \in (\mathbb{B}_T)^h$ platí $D^h \in \mathbb{B}^h = \mathcal{B}(X^h, Y^h, I^h)$. Stačí tedy ukázat, že $D^h \models T^h$. To plyne z $D^h \in (\mathbb{B}_T)^h$, protože pak $D \models T$, což znamená, že $D^{h'} \models T^{h'}$ pro libovolné $h' : L \to 2$, speciálně tedy $D^h \models T^h$. □

Opačná inkluze v (2) obecně neplatí. Jako protipříklad vezměme podmíněný konceptuální svaz $\mathbb{B}^1$ z příkladu 41 a uvažujme $T = \{y_2 \sqsubseteq y_1\}$. Při realizaci podle totální reality h_3 platí:

$$(\mathbb{B}^1_T)^{h_3} \not\supseteq \mathcal{B}_{T^{h_3}}(X^{h_3}, Y^{h_3}, I^{h_3}).$$

Všimněme si, že pro podmíněný konceptuální svaz $\mathbb{B}^2$ ze stejného příkladu, stejnou množinu T a libovolnou totální realitu h_i platí:

$$(\mathbb{B}^2_T)^{h_i} = \mathcal{B}_{T^{h_i}}(X^{h_i}, Y^{h_i}, I^{h_i}).$$

Definice 48 (stabilní podmíněný konceptuální svaz)

L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X,Y,I)$, který pro danou množinu T AD formulí nad $(Y, \approx_Y)$ a pro každou totální realitu $h : L \to 2$ splňuje:

$$(\mathbb{B}_T)^h = \mathcal{B}_{T^h}(X^h, Y^h, I^h),$$

nazveme T-*stabilní*.

Není-li potřeba množinu T AD formulí explicitně zmiňovat, budu namísto „T-stabilní" psát jen „stabilní".

Stabilita je téměř nezbytná vlastnost podmíněného konceptuálního svazu, který má být omezován AD formulemi. Jedině T-stabilní podmíněné konceptuální svazy totiž po omezení AD formulemi z množiny T zaručeně obsahují podmíněný koncept pro každý formální koncept ve všech příslušných omezených konceptuálních svazech. Při volbě nestabilního podmíněného konceptuálního svazu $\mathbb{B}$ tak uživatel vždy podstupuje riziko, že nějaký potenciálně důležitý formální koncept nebude v $\mathbb{B}$ po omezení AD formulemi reprezentován žádným podmíněným konceptem.

Věta 49

$\mathbb{B}^*(X,Y,I)$ je T-stabilní L-podmíněný konceptuální svaz pro libovolnou množinu T AD formulí nad $(Y, \approx_Y)$.

Důkaz

Označme $\mathbb{B}^*(X,Y,I) = \mathbb{B}^*$. To, že $\mathbb{B}^*$ je L-podmíněný konceptuální svaz, bylo ukázáno v [2] (Theorem 10, strana 26). Zde ukážu, že je stabilní. Uvažujme libovolnou množinu T AD formulí nad $(Y, \approx_Y)$. Stačí ukázat, že pro libovolnou totální realitu $h_i : L \to 2$ platí:

$$(\mathbb{B}^*_T)^{h_i} \supseteq \mathcal{B}_{T^{h_i}}(X^{h_i}, Y^{h_i}, I^{h_i}),$$

protože opačná inkluze platí vždy (Věta 47).

Ať tedy $C^i_{x_i}$ je formální koncept z $\mathcal{B}_{T^{h_i}}(X^{h_i}, Y^{h_i}, I^{h_i})$. Podle Věty 42 můžeme všechny podmíněné koncepty z $\mathbb{B}^*$ charakterizovat n-ticemi formálních konceptů $(C^1_{x_1}, \ldots, C^n_{x_n}) \in \mathcal{B}^1 \times \ldots \times \mathcal{B}^n$. Uvažme n-tici:

$$\overline{C} = (C^1_0, \ldots, C^{i-1}_0, C^i_{x_i}, C^{i+1}_0, \ldots, C^n_0),$$

kde C^j_0 označuje vždy ten nejmenší formální koncept v $\mathcal{B}(X^{h_j}, Y^{h_j}, I^{h_j})$, a příslušný podmíněný koncept v $\mathbb{B}^*$ označme $f(\overline{C})$.

Pro každý formální koncept $C^j_{x_j}$ z $\overline{C}$ platí $C^j_{x_j} \models T^{h_j}$ ($C^i_{x_i} \models T^{h_i}$ platí z předpokladu, všechny ostatní jistě splňují libovolnou T). Pak platí $f(\overline{C}) \models T$ a tedy $f(\overline{C}) \in \mathbb{B}^*_T$. Podle Věty 42 pak $(f(\overline{C}))^{h_i} = C^i_{x_i} \in (\mathbb{B}^*_T)^{h_i}$.

□

5.1.1 Dolní hranice pro počet prvků stabilních podmíněných konceptuálních svazů

Označme $\mathcal{B}^i = \mathcal{B}(X^{h_i}, Y^{h_i}, I^{h_i})$. Z Věty 42 plyne $|\mathbb{B}^*(X,Y,I)| = |\mathcal{B}^1 \times \ldots \times \mathcal{B}^n|$, horní hranice pro počet prvků libovolného L-podmíněného konceptuálního svazu $\mathbb{B} \subseteq \mathbb{B}^*(X,Y,I)$ je tedy rovna součinu mohutností $|\mathcal{B}^i|$ konceptuálních svazů $\mathcal{B}^i$ pro všechny totální reality $h_i : L \to 2$.

Co se týče dolní hranice, $\mathbb{B}$ musí pro všechna h_i splňovat $\mathbb{B}^{h_i} = \mathcal{B}^i$, $\mathbb{B}$ tak musí obsahovat přinejmenším stejný počet prvků, jako má nejvíce početný konceptuální svaz $\mathcal{B}^i$ (pro každé $C^i_j \in \mathcal{B}^i$ musí existovat alespoň jedno $D \in \mathbb{B}$ takové, že $D^{h_i} = C^i_j$). Celkově tedy pro obecný podmíněný konceptuální svaz $\mathbb{B}$ platí:

$$\max_i(|\mathcal{B}^i|) \leq |\mathbb{B}| \leq |\mathcal{B}^1 \times \ldots \times \mathcal{B}^n|.$$

V této podkapitole zpřesním dolní hranici pro podmíněné konceptuální svazy, které jsou T-stabilní pro danou množinu AD formulí T (Věta 51). Nejprve uvedu jedno pomocné lemma.

Lemma 50

Pro každý L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X,Y,I)$, každou množinu T AD formulí nad $(Y, \approx_Y)$ a každou totální realitu $h_i : L \to 2$ platí:

$$|\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}}| \leq |\mathbb{B} \setminus \mathbb{B}_T|.$$

Je-li $\mathbb{B}$ T-stabilní, pak navíc platí:

$$|\mathcal{B}^i_{T^{h_i}}| \leq |\mathbb{B}_T|.$$

Důkaz

Pro každé $C^i_j \in (\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}})$ existuje alespoň jedno $D \in \mathbb{B}$ takové, že $D^{h_i} = C^i_j$. Protože C^i_j nesplňuje T^{h_i} ($C^i_j \notin \mathcal{B}^i_{T^{h_i}}$), platí, že ani žádné z takových D nesplňuje T, a tak všechna taková D leží v $(\mathbb{B} \setminus \mathbb{B}_T)$. Jinými slovy: pro každý formální koncept z $(\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}})$ existuje alespoň jeden podmíněný koncept v $(\mathbb{B} \setminus \mathbb{B}_T)$.

Kdyby platilo $|\mathbb{B}_T| < |\mathcal{B}^i_{T^{h_i}}|$, existovalo by $C^i_j \in \mathcal{B}^i_{T^{h_i}}$, pro které neexistuje $D \in \mathbb{B}_T$ takové, že $D^{h_i} = C^i_j$. Platilo by tedy $(\mathbb{B}_T)^{h_i} \not\supseteq \mathcal{B}^i_{T^{h_i}}$, což znamená, že $\mathbb{B}$ by nebyl T-stabilní. □

Věta 51

Ať T je množina AD formulí nad $(Y, \approx_Y)$. Pro libovolný T-stabilní podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ platí:

$$\max_i(|\mathcal{B}^i_{T^{h_i}}|) + \max_i(|\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}}|) \leq |\mathbb{B}|.$$

Důkaz

Předpokládejme, že $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ je T-stabilní podmíněný konceptuální svaz a přitom platí:

$$|\mathbb{B}| = |\mathbb{B}_T| + |\mathbb{B} \setminus \mathbb{B}_T| < \max_i(|\mathcal{B}^i_{T^{h_i}}|) + \max_i(|\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}}|). \tag{3}$$

Protože $\mathbb{B}$ je T-stabilní, platí, podle Lemmatu 50, $|\mathcal{B}^i_{T^{h_i}}| \leq |\mathbb{B}_T|$ pro všechna i, tedy i pro to i, pro které nastává ono maximum. Nerovnost (3) tak můžeme přepsat následovně:

$$|\mathbb{B} \setminus \mathbb{B}_T| < \max_i(|\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}}|).$$

To ovšem znamená, že existuje i takové, že $|\mathbb{B} \setminus \mathbb{B}_T| < |\mathcal{B}^i \setminus \mathcal{B}^i_{T^{h_i}}|$ a to je, opět podle Lemmatu 50, spor s tím, že $\mathbb{B}$ je podmíněný konceptuální svaz. □

5.1.2 Algoritmus STABILIZE

V této podkapitole uvedu pseudokód algoritmu, který k danému podmíněnému konceptuálnímu svazu $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ a dané množině T AD formulí nad $(Y, \approx_Y)$ sestaví T-stabilní podmíněný konceptuální svaz $\mathbb{B}' \supseteq \mathbb{B}$, přičemž počet prvků výsledného $\mathbb{B}'$ bude nanejvýš dvojnásobný oproti počtu prvků $\mathbb{B}$.

V následujících pseudokódech označuje H množinu všech totálních realit $h_1, \ldots, h_n : L \to 2$ a a_i označuje atom z L příslušný totální realitě h_i.

Algoritmus STABILIZE
Vstupem je L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ a množina T AD formulí nad $(Y, \approx_Y)$

1. $\mathbb{B}' \leftarrow \mathbb{B}$
2. pro všechna $D \in \mathbb{B}$ opakuj
   (a) $H' \leftarrow \{h \in H \mid D^h \models T^h\}$

(b) jestliže $H' = \emptyset$ nebo $H' = H$, pak

i. přejdi na další D

(c) jinak

i. $D' \leftarrow$ STABILIZE-CONCEPT$(D,\, H')$

ii. $\mathbb{B}' \leftarrow \mathbb{B}' \cup \{D'\}$

3. vrať $\mathbb{B}'$

Algoritmus STABILIZE-CONCEPT
Vstupem je L-podmíněný koncept $D \in \mathbb{B}^*(X, Y, I)$ a množina realit $H' \subseteq H$

1. alokuj pole pro n prvků a ulož jej do proměnné N
2. pro $i \leftarrow 0$ do $n - 1$ opakuj

   (a) jestliže $h_{i+1} \in H'$, pak

   i. $N[i] \leftarrow D^{h_{i+1}}$

   (b) jinak

   i. $N[i] \leftarrow ((Y^{h_{i+1}})^{\downarrow}, Y^{h_{i+1}})$

3. sestav podmíněné množiny A v X, B v Y nálsedovně

   (a) $A(x) \leftarrow \bigvee\{a_i \mid x \in Ext(N[i])\}$

   (b) $B(y) \leftarrow \bigvee\{a_i \mid y \in Int(N[i])\}$

4. vrať (A, B)

Algoritmus je založený na následujícím pozorování. Je-li daný L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$ vůči dané množině T AD formulí nad $(Y, \approx_Y)$ nestabilní, pak existuje L-podmíněný koncept $D \in \mathbb{B}$ a totální reality $h_i, h_j : L \to 2$ tak, že platí:

$$D^{h_i} \models T^{h_i} \text{ a přitom } D^{h_j} \not\models T^{h_j}.$$

Pokud pak pro každé takové D přidáme do $\mathbb{B}$ L-podmíněný koncept D', který pro každou totální realitu $h : L \to 2$, pro kterou platí $D^h \models T^h$, splňuje:

$$D' \models T \text{ a } (D')^h = D^h,$$

pak $\mathbb{B}$ bude T-stabilní. Správnost algoritmu plyne z tohoto pozorování a z Věty 42.

5.2 Zobecnění výsledků z klasické FCA

V poslední kapitole uvádím varianty některých výsledků z klasické FCA (konkrétně Vět 11, 12, 13 a 14) v podmíněné FCA za předpokladu, že uživatel použije stabilní podmíněný konceptuální svaz.

Po zbytek kapitoly uvažujme množinu T AD formulí nad $(Y, \approx_Y)$ a T-stabilní L-podmíněný konceptuální svaz $\mathbb{B} \subseteq \mathbb{B}^*(X, Y, I)$.

Věta 52

Pro každou totální realitu $h : L \to 2$ je $(\mathbb{B}_T)^h$ zdola omezený.

Důkaz

Pro každou totální realitu $h : L \to 2$ platí $(\mathbb{B}_T)^h = \mathcal{B}_{T^h}(X^h, Y^h, I^h)$, což je zdola omezená množina pro libovolné T^h podle Věty 11. □

V následujícím lemmatu označuje $X^{\uparrow}$ podmíněnou množinu v $(Y, \approx_Y)$ (kde X chápeme jako podmíněnou množinu v $(X, \approx_X)$). Dále pro $B = \{y_i, \dots, y_j\} \subseteq Y$ a realitu $h : L \to K$ označuje B^h množinu $\{y_i^h, \dots, y_j^h\}$.

Lemma 53

Pokud T neobsahuje AD formuli $A \sqsubseteq B$, takovou, že pro nějaké $y \in A$ platí:

$$X^{\uparrow}(y) > \bigvee_{y' \in B} X^{\uparrow}(y'), \tag{4}$$

pak pro každou totální realitu $h : L \to 2$ a každou AD formuli $A' \sqsubseteq B' \in T$ platí, že neexistuje $y^h \in (A')^h$ takové, že:

$$y^h \in (X^h)^{\uparrow} \text{ a zároveň } (y')^h \notin (X^h)^{\uparrow} \text{ pro každé } (y')^h \in (B')^h. \tag{5}$$

Důkaz

Kdyby existovala totální realita $h : L \to 2$, AD formule $A' \sqsubseteq B' \in T$ a $y^h \in (A')^h$ tak, že platí (5), platilo by:

$$h(X^{\uparrow}(y)) = 1 > 0 = h(\bigvee_{y' \in B} X^{\uparrow}(y')),$$

a tedy $A' \sqsubseteq B'$ by splňovala (4). □

Věta 54

Pokud T neobsahuje AD formuli $A \sqsubseteq B$, pro kterou platí (4), pak pro každou totální realitu $h : L \to 2$ platí, že $(\mathbb{B}_T)^h$ je shora omezený.

Důkaz

Pro každou totální realitu $h : L \to 2$ množina T^h podle Lemmatu 53 neobsahuje AD formuli $A^h \sqsubseteq B^h$, pro kterou by platilo, že A^h obsahuje nějaký atribut sdílený všemi objekty a zároveň v B^h žádný takový atribut neleží, a tak podle Věty 12 je $\mathcal{B}_{T^h}(X^h, Y^h, I^h) = (\mathbb{B}_T)^h$ shora omezený. □

Věta 55

Pokud jsou všechny AD formule z T tvaru $\{y\} \sqsubseteq \{y'\}$ pro nějaká $y, y' \in Y$, pak pro každou totální realitu $h : L \to 2$ platí, že $(\mathbb{B}_T)^h$ je úplný svaz, který je navíc $\bigvee$-podsvazem množiny $\mathcal{B}(X^h, Y^h, I^h)$.

Důkaz

Pro každou totální realitu $h : L \to 2$ bude T^h obsahovat pouze AD formule tvaru $\{y^h\} \sqsubseteq \{(y')^h\}$, a tak podle Věty 13 bude požadované vlastnosti splňovat $\mathcal{B}_{T^h}(X^h, Y^h, I^h) = (\mathbb{B}_T)^h$. □

V následujícím lemmatu je řeč o disjunktních atributech. V podmíněném kontextu (X, Y, I) atributy y, $y' \in Y$ nazývám disjunktní, když pro všechna $x \in X$ platí $I(x, y) \wedge I(x, y') = 0$. V nepodmíněném formálním kontextu (X', Y', I') pak atributy y, $y' \in Y'$ nazývám disjunktní, když neexistuje objekt $x \in X$, pro který by platilo $(x, y) \in I'$ a zároveň $(x, y') \in I'$

Lemma 56

Pro disjunktní atributy y, $y' \in Y$ a každou totální realitu $h : L \to 2$ platí, že i atributy y^h, $(y')^h \in Y^h$ jsou disjunktní.

Důkaz

Předpokládejme, že y, $y' \in Y$ jsou disjunktní a přitom existuje totální realita $h : L \to 2$, podle které y^h, $(y')^h \in Y^h$ disjunktní nejsou. Pak tedy existuje $x^h \in X^h$ takové, že $I^h(x^h, y^h) = I^h(x^h, (y')^h) = 1$. Protože I je extensionální, platí $I^h(x^h, y^h) = h(I(x, y))$, $I^h(x^h, (y')^h) = h(I(x, y'))$. Dále protože y, $y' \in Y$ jsou disjunktní, platí $I(x, y) \wedge I(x, y') = 0$. Pak ale

$$h(I(x, y)) \wedge h(I(x, y')) = 1 \wedge 1 \neq 0 = h(0) = h(I(x, y) \wedge I(x, y'))$$

a to je spor s tím, že h je homomorfismus BA. □

Věta 57

Pokud T obsahuje pro každou dvojici disjunktních atributů y, $y' \in Y$ AD formuli $A \sqsubseteq B$, pro kterou platí buď:

$$y \in A,\ y' \in B \text{ a všechny atributy z } B \text{ jsou disjunktní, nebo} \tag{6}$$

$$y' \in A,\ y \in B \text{ a všechny atributy z } B \text{ jsou disjunktní,} \tag{7}$$

pak pro každou totální realitu $h : L \to 2$ platí, že $(\mathbb{B}_T)^h$ je strom.

Důkaz

Pro každou totální realitu $h : L \to 2$ a každou dvojici nedisjunktních atributů y^h, $(y')^h \in Y^h$ bude T^h obsahovat AD formuli, která splňuje (6) nebo (7). Plyne to z Lemmatu 56: disjunktnost atributů na pravých stranách AD formulí zůstane po realizaci zachována a kromě toho nemůže existovat dvojice nedisjunktních atributů y^h, $(y')^h \in Y^h$, které by před realizací byly disjunktní, a tak žádaná AD formule opravdu existuje pro každou takovou dvojici. A tak podle Věty 14 bude $\mathcal{B}_{T^h}(X^h, Y^h, I^h) = (\mathbb{B}_T)^h$ strom. $\square$

Závěr

Formální konceptuální analýza ve své základní podobě obnáší problém, kterým je potenciálně obrovská velikost konceptuálního svazu. Tento problém lze řešit například za pomoci AD formulí. Dalším problémem, se kterým se uživatel FCA (a nejen uživatel FCA) může potýkat, je neúplnost dat; tu je pak v některých případech možné řešit přechodem k podmíněné variantě FCA.

Práce měla za cíl implementovat AD formule v podmíněné FCA. Přestože definice základních pojmů je přímočará, vyvstal při implementaci zajímavý problém: některé podmíněné konceptuální svazy, které by za jiných okolností mohly být bez problému použity, jsou naprosto nevhodné, má-li uživatel v plánu je omezovat AD formulemi. Stabilita, ona vlastnost, která v tomto smyslu rozlišuje mezi vhodnými a nevhodnými podmíněnými konceptuálními svazy, je jasně definovatelná. Existuje algoritmus, který libovolný podmíněný konceptuální svaz učiní stabilním vůči dané množině AD formulí a nijak zásadně přitom nenavýší počet jeho prvků. Navíc lze za předpokladu stability podmíněného konceptuálního svazu lehce zobecnit a využívat některé důležité výsledky z teorie AD formulí v klasické variantě FCA.

Conclusions

Formal concept analysis in its basic form involves a problem, which is the potentially enormous size of the concept lattice. This problem can be addressed, for instance, by usage of AD formulas. Another issue that FCA users (and not only FCA users) may encounter is the incompleteness of data; that may be, in some cases, resolved by transitioning to the conditional variant of FCA.

The goal of the work was to implement AD formulas in conditional FCA. Although the basic definitions are straightforward, an interesting problem arose during implementation: some conditional concept lattices, which could be easily used under different circumstances, are entirely unsuitable if the user plans to restrict them by AD formulas. Stability, the property that distinguishes between suitable and unsuitable conditional concept lattices in this sense, is precisely definable. An algorithm exists that makes any conditional concept lattice stable with respect to a given set of AD formulas, without significantly increasing its number of elements. Moreover, assuming the stability of the conditional concept lattice, it is easy to generalize and utilize some important results from the theory of AD formulas in the basic form of FCA.

A Obsah elektronických dat

`text/`

Adresář s textem práce ve formátu PDF, vytvořený s použitím závazného stylu KI PřF UP v Olomouci pro závěrečné práce, a ZIP archiv se všemi soubory potřebnými pro bezproblémové vygenerování PDF dokumentu textu.

Literatura

[1] Bělohlávek R. (2008): *Introduction to Formal Concept Analysis.*

[2] Krupka M., Laštovička J. (2017): *Incomplete information and concept lattices.* International Journal of General Systems.

[3] Bělohlávek R., Sklenář V. (2005): *Formal Concept Analysis Constrained by Attribute-Dependency Formulas.* Proc. ICFCA 2005 (Lecture Notes in Computer Science), vol. 3403. New York: Springer-Verlag, pp. 176–191.

[4] Bělohlávek R., Vychodil V. (2009): *Formal Concept Analysis With Background Knowledge: Attribute Priorities.* IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).

[5] Givant S., Halmos P. (2009): *Introduction to Boolean Algebras.* New York: Springer-Verlag.