Determinacja zapytań koniunkcyjnych w semantyce multizbiorowej

(Determinacy of conjunctive queries under bag semantics)

Jarosław Kwiecień

Praca magisterska

Promotor: prof. dr hab. Jerzy Marcinkowski

Uniwersytet Wrocławski
Wydział Matematyki i Informatyki
Instytut Informatyki

29 czerwca 2022

Streszczenie

Problem determinacji zapytań jest jednym z podstawowych zagadnień teorii baz danych. Przez ostatnie dwie dekady ukazało się wiele prac opisujących złożoność obliczeniową (głównie (nie)rozstrzygalność) tego problemu dla różnych klas zapytań. Wszystkie one dotyczyły jednak semantyki zbiorowej, tj. sytuacji, w której znaczenie ma jedynie fakt znalezienia odpowiedzi na zapytanie, a nie liczby sposobów, na jakie można tę odpowiedź znaleźć. Jednak to ten drugi przypadek, nazywany semantyką multizbiorową, jest bliższy praktycznym zastosowaniom. W niniejszej pracy prezentujemy nowe wyniki dotyczące złożoności obliczeniowej problemu determinacji zapytań koniunkcyjnych, ścieżkowych oraz unii zapytań koniunkcyjnych w semantyce multizbiorowej.

The problem of query determinacy is one of the fundamental aspects of the database theory. Over the last two decades there have been a lot of papers describing the computational complexity (mostly (un)decidability) of this problem for various classes of queries. However, they all concerned set semantics, i.e. a situation in which only the fact of finding an answer to a query matters, not the number of ways in which that answer can be found. However, it is the latter case, called multiset (or bag) semantics, that is closer to practical applications. In this paper, we present new results on the computational complexity of the determinacy problem for conjunctive queries, path queries and unions of conjunctive queries under multi-set semantics.

Spis treści

1 Wprowadzenie **7**

1.1 Wstęp . . . 7

1.2 Preliminaria . . . 8

2 Zapytania ścieżkowe **15**

2.1 Dowód implikacji $\Leftarrow$ . . . 16

2.1.1 q-spacery i ich redukcje . . . 16

2.1.2 Bazy danych a przekształcenia liniowe . . . 18

2.1.3 Wykorzystanie dwóch redukcji q-spacerów . . . 19

2.2 Dowód implikacji $\Rightarrow$ . . . 20

3 Unie zapytań koniunkcyjnych (UCQ) **23**

4 Boole'owskie zapytania koniunkcyjne **29**

4.1 Warunek równoważny . . . 29

4.2 Dowód lematu 4.8 ($\Leftarrow$) . . . 31

4.3 Dowód lematu 4.8 ($\Rightarrow$) . . . 32

4.3.1 Podstawowe narzędzia . . . 32

4.3.2 Poszukiwanie dobrego S - dowód lematu 4.15 . . . 34

4.3.3 Wskazanie kontrprzykładu - dowód lematu 4.16 . . . 35

4.4 Uwaga o homomorfizmach struktur . . . 39

4.4.1 Gęstości homomorfizmów . . . 40

5 Pozostałe przypadki CQ - otwarte problemy **43**

5.1 Unarne zapytania koniunkcyjne . . . 43

Bibliografia **47**

Rozdział 1

Wprowadzenie

1.1 Wstęp

Zapewne każdy z czytelników tej pracy słyszał o bazach danych. Są wszechobecne i przechowują informacje o niemal każdym aspekcie naszego życia - od znajomości i zainteresowań po stan zdrowia, czy finansów. Dostęp do tych informacji realizowany jest przez *zapytania* sformułowane w pewnym języku formalnym, dla których silnik bazodanowy zwraca zbiór (lub multizbiór) krotek składających się z elementów bazy spełniających zadane zapytanie.

Ze względu na swoją kluczową rolę w funkcjonowaniu współczesnych systemów[1], bazy danych są przedmiotem intensywnych badań od dziesięcioleci. Jednym z naturalnych problemów pojawiających się w tej tematyce jest: „Czy znając odpowiedzi na zapytania $v_1, v_2, ..., v_n$ mam już wszystkie informacje potrzebne, by przewidzieć odpowiedź na zapytanie q”? Nietrudno wyobrazić sobie zastosowania tego problemu w takich dziedzinach, jak bezpieczeństwo informacji czy optymalizacja silników bazodanowych.

Problem ten, nazwany *problemem determinacji zapytań* był poruszany w wielu pracach w różnych wariantach - w zależności od przyjętych założeń (przede wszystkim tych dotyczących klasy rozważanych zapytań) dowodzono jego rozstrzygalności ([15], [2]) lub nierozstrzygalności ([16], [9], [10]). W wyniku tych prac, powstała dość dokładna klasyfikacja, podsumowana w pracy [14], gdzie przedstawione są eleganckie techniki badania determinacji przy użyciu takich narzędzi teorii baz danych, jak *Tuple Generating Dependencies* i *Chase*.

Wszystkie te odkrycia działają jednak przy założeniu, że odpowiedzią na zapytania jest pewien *zbiór* krotek (mówimy wtedy o *semantyce zbiorowej*, ang. *set semantics*). Jednak modelem lepiej opisującym działanie prawdziwych baz danych jest taki, w którym dla zapytań zwracane są *multizbiory* krotek (semantyka mul-

[1] z pewnością też ze względu na prosty, elegancki model matematyczny, który je opisuje

tizbiorowa, ang. *bag semantics*). Na tą lukę pomiędzy teorią a praktyką zwrócono uwagę w pracy [6] wykazując też, że techniki dowodzenia używane w semantyce zbiorowej nie przenoszą się na semantykę multizbiorową. Co więcej, okazuje się, że problemy dotyczące semantyki multizbiorowej są znacznie trudniejsze i wiele z nich wciąż pozostaje otwartych. Flagowym przykładem jest tu *problem zawierania zapytań koniunkcyjnych* - podstawowy problem związany z optymalizacją zapytań. Już w latach 70-tych udowodniono, że w semantyce zbiorowej jest to problem NP-zupełny. Z kolei w semantyce multizbiorowej wciąż nie wiadomo, czy jest on rozstrzygalny, a jedna z ostatnich prac [1] pokazała jego równoważność z pewnym problemem z teorii informacji, którego rozstrzygalność również jest otwartym pytaniem od wielu lat.

W niniejszej pracy opisujemy podjętą przez nas pierwszą znaną nam próbę zbadania rozstrzygalności problemu determinacji w semantyce multizbiorowej. Wykażemy nierozstrzygalność tego problemu dla unii zapytań koniunkcyjnych, a także jego rozstrzygalność w przypadku zapytań ścieżkowych oraz boole'owskich zapytań koniunkcyjnych. Ten ostatni przypadek jest szczególnie ciekawy, ponieważ okazuje się być bezpośrednio związany z zagadnieniem teorii grafów dotyczącej zliczania homomorfizmów. Nasz wynik dowodzi tezy postawionej w pracy [7] z 1979 roku, a później przytaczanej wielokrotnie w kolejnych publikacjach Lovasza ([13], [4]), która mówiła, że dla dla zbioru parami nieizomorficznych grafów, funkcje zliczające homomorfizmy z tych grafów w pewien zadany graf są od siebie niezależne. Dokładnie opisujemy to w rozdziale 4

W ostatnim rozdziale przedstawiamy nierozwiązane przez nas problemy dotyczące determinacji zapytań w semantyce multizbiorowej i opisujemy nasze przemyślenia dotyczące przypadku zapytań unarnych, wskazując potencjalne kierunki rozwoju w tym temacie.

Wyniki przedstawione w tej pracy są efektem wspólnych działań autora z prof. Jerzym Marcinkowskim oraz Piotrem Ostropolskim-Nalewają. Zostały one również opublikowane w artykule [11], który zostanie zaprezentowany podczas międzynarodowej konferencji *ACM SIGMOD/PODS 2022*.

1.2 Preliminaria

Konwencje

Dla multizbioru A oraz elementu x napis $A[x]$ oznacza krotność występowania x w A. W szczególności $A[x] > 0 \iff x \in A$. Gdy piszemy $\bar{x}$, zawsze mamy na myśli krotkę (której liczba elementów zazwyczaj wynika z kontekstu). Np. pisząc $\varphi(\bar{x})$ mamy na myśli formułę/funkcję pewnej arności, niekoniecznie równej 1 ($\bar{x}$ jest tu krotką zmiennych). $\mathbb{R}_{\geq 0}$ oznacza zbiór nieujemnych liczb rzeczywistych (analogicznie dla $\mathbb{Q}, \mathbb{N}$ itd.). Przyjmujemy, że $0^0 = 1$.

Sygnatura

W niniejszej pracy będziemy rozważać jedynie skończone relacyjne sygnatury (to znaczy skończone zbiory symboli relacyjnych dowolnej arności), z wyjątkiem miejsc w których zaznaczono, że jest inaczej.

Baza danych

Niech $\Sigma = \{R_1, R_2, ..., R_n\}$ będzie sygnaturą. Wtedy *bazą danych* nad sygnaturą Σ nazywamy dowolną skończoną strukturę nad tą sygnaturą, to znaczy strukturę postaci $\mathcal{D} = \langle D, R_1^{\mathcal{D}}, R_2^{\mathcal{D}}, ..., R_n^{\mathcal{D}} \rangle$, gdzie D jest skończonym zbiorem - uniwersum struktury $\mathcal{D}$ (oznaczanym również jako $|\mathcal{D}|$), a $R_1^{\mathcal{D}}, ..., R_n^{\mathcal{D}}$ są relacjami nad zbiorem D - o arnościach odpowiadających arnościom symboli $R_1, ..., R_n$. Czasem, gdy nie będzie to prowadziło do niejednoznaczności, będziemy utożsamiać bazę danych $\mathcal{D}$ z jej uniwersum $|\mathcal{D}|$ (pisząc na przykład „$x \in \mathcal{D}$"). W niniejszej pracy wszystkie rozważane struktury będą skończone, zatem określenia *struktura* i *baza danych* są traktowane jak synonimy.

Dla zdania logicznego φ oznaczenie $\mathcal{D} \models \varphi$ mówi „φ jest prawdziwe w bazie danych $\mathcal{D}$". W szczególności, dla atomu $R(\bar{x})$, $\mathcal{D} \models R(\bar{a}) \iff \bar{a} \in R^{\mathcal{D}}$.

Homomorfizm

Niech $\mathcal{D}_1, \mathcal{D}_2$ będą bazami danych nad sygnaturą Σ. Odwzorowanie $f : |\mathcal{D}_1| \to |\mathcal{D}_2|$ jest homomorfizmem struktur $\mathcal{D}_1$ i $\mathcal{D}_2$, gdy zachodzi

$$\forall R \in \Sigma \ \forall \bar{x} \in R^{\mathcal{D}_1} \ f(\bar{x}) \in R^{\mathcal{D}_2}$$

gdzie dla $\bar{x} = \langle x_1, ..., x_k \rangle$ przez $f(\bar{x})$ rozumiemy krotkę $\langle f(x_1), ..., f(x_k) \rangle$.

Zbiór wszystkich homomorfizmów $\mathcal{D}_1 \to \mathcal{D}_2$ oznaczamy przez $hom(\mathcal{D}_1, \mathcal{D}_2)$.

Zapytanie koniunkcyjne

Zapytaniami koniunkcyjnymi (ang. Conjunctive Query, CQ), dla pewnej sygnatury Σ, nazywamy formuły logiki pierwszego rzędu postaci $\varphi(\bar{x}) = \exists \bar{y} \ \psi(\bar{x}, \bar{y})$, gdzie $\psi(\bar{x}, \bar{y})$ jest koniunkcją atomów. $\bar{x} = \langle x_1, ..., x_k \rangle$ nazywamy *zmiennymi wolnymi* zapytania, a $\bar{y} = \langle y_1, ..., y_m \rangle$ - *zmiennymi skwantyfikowanymi*. Liczba zmiennych wolnych wyznacza *arność* zapytania. Np. gdy $\bar{x} = \langle x_1, x_2 \rangle$, to zapytanie φ jest binarne.

Z każdym zapytaniem koniunkcyjnym $\varphi(\bar{x}) = \exists \bar{y} \ \psi(\bar{x}, \bar{y})$ związana jest pewna struktura F_φ, nazywana *frozen body*, zdefiniowana w następujący sposób: uniwersum F_φ stanowi zbiór zmiennych wystepujących w krotkach $\bar{x}, \bar{y}$. Ponadto, dla dowolnego

n, dowolnej $n-$arnej relacji R oraz krotki $\bar{z} \in |F_\varphi|^n$,

$$F_\varphi \models R(\bar{z}) \iff R(\bar{z}) \in \psi$$

Sformułowanie $R(\bar{z}) \in \psi$ należy tu rozumieć jako „$R(\bar{z})$ jest jednym ze składników koniunkcji ψ"

Dla zapytania koniunkcyjnego φ zdefiniowanego jak wyżej, *wynikiem zapytania w bazie* $\mathcal{D}$ jest **multizbiór** $\varphi(\mathcal{D})$ zdefiniowany następująco:

$$\varphi(\mathcal{D})[\bar{x}] = |\{\bar{y} \mid \mathcal{D} \models \psi(\bar{x}, \bar{y})\}|$$

W wielu pracach dotyczących baz danych $\varphi(\mathcal{D})$ byłby zbiorem, a nie multizbiorem. W takiej sytuacji mówimy o *semantyce zbiorowej.*

Problemy związane z zapytaniami koniunkcyjnymi mogą być rozważane w obu tych wersjach, przy czym przeważnie wersja *zbiorowa* jest znacznie prostsza od *multizbiorowej.* Dlatego my zajmujemy się semantyką *multizbiorową.*

Zapytanie ścieżkowe

Zapytanie ścieżkowe (ang. Path Query, PQ), nad binarną sygnaturą Σ, to binarne zapytania koniunkcyjne postaci

$$\varphi(x_1, x_2) = \exists y_1, ..., y_n\ R_1(x_1, y_1) \wedge R_2(y_1, y_2) \wedge ... \wedge R_n(y_{n-1}, y_n) \wedge R_{n+1}(y_n, x_2)$$

Innymi słowy, jest to zapytanie „Ile jest ścieżek z x_1 do x_2 składających się z $n+1$ krawędzi o etykietach, kolejno: $R_1, R_2, ..., R_{n+1}$". Takie zapytania będziemy identyfikować ze słowami nad alfabetem Σ. Np. zapytanie napisane powyżej identyfikujemy ze słowem $R_1...R_{n+1}$. Szczególny przypadek stanowi tu słowo puste, z którym identyfikujemy zapytanie $\varphi(x_1, x_2) =$ "$x_1 = x_2$", choć zgodnie z definicją powyżej nie jest to zapytanie ściezkowe.

Zapytanie koniunkcyjne boole'owskie

Zapytanie boole'owskie (ang. Boolean Conjunctive Query, BCQ) to zapytanie koniunkcyjne bez zmiennych wolnych. Takie zapytanie będziemy identyfikować z jego *frozen body* pisząc po prostu φ zamiast F_φ tam, gdzie z kontekstu będzie wynikało, że chodzi o strukturę (np. gdy będziemy mówili o homomorfizmach).

Dla zapytania boole'owskiego q, jego wynik w strukturze $\mathcal{D}$ to multizbiór zawierający pustą krotkę pewną skończoną liczbę razy. Dokładniej:

Obserwacja 1.1. $q(\mathcal{D})[\langle\rangle] = |hom(q, \mathcal{D})|$

Dla uproszczenia będziemy pisać $q(\mathcal{D})$ zamiast $q(\mathcal{D})[\langle\rangle]$.

Unią boole'owskich zapytań koniunkcyjnych nazywamy logiczną alternatywę skończonej liczby boole'owskich zapytań koniunkcyjnch. Wynikiem zapytania postaci $Q = q_1 \vee q_2 \vee ... \vee q_n$ jest $Q(\mathcal{D}) = \sum_{i=1}^{n} q_i(\mathcal{D})$.

Determinacja zapytań

Niech $V = \{v_1, ..., v_n\}$ będzie pewnym skończonym zbiorem zapytań (w tym przypadku: zapytań koniunkcyjnch lub unii zapytań koniunkcyjnych). Będziemy nazywać go zbiorem *widoków*. Także niech q będzie zapytaniem. Mówimy, że V *determinje* q, gdy dla każdej pary baz danych $\mathcal{D}, \mathcal{D}'$ zachodzi następująca implikacja:

$$(\forall v \in V \ v(\mathcal{D}) = v(\mathcal{D}')) \Rightarrow q(D) = q(D')$$

Oznaczenie: $V \xrightarrow{\text{bag}} q$ lub $V \xrightarrow{\text{set}} q$, w zależności od tego, czy napisy $v(D), q(D)$ traktujemy jako zbiory, czy multizbiory.

Obserwacja 1.2. *Warunek* $V \xrightarrow{bag} q$ *jest równoważny następującemu warunkowi:*

Istnieje funkcja f taka, że dla każdej bazy danych $\mathcal{D}$ zachodzi

$$q(\mathcal{D}) = f(v_1(\mathcal{D}), v_2(\mathcal{D}), ..., v_n(\mathcal{D}))$$

Zawieranie zapytań koniunkcyjnych

Zapytanie q zawiera się w zapytaniu q' ($q \subseteq q'$), gdy dla każdej bazy danych $\mathcal{D}$, $q(\mathcal{D}) \subseteq q'(\mathcal{D})$. Zawieranie $q(\mathcal{D}) \subseteq q'(\mathcal{D})$ możemy traktować jako zawieranie zbiorów lub multizbiorów[2]. W tej pracy będziemy, o dziwo, będziemy odwoływać się do zawierania zapytań w wersji zbiorowej, dla jasności oznaczając je jako $q \subseteq_{\text{set}} q'$.

Fakt 1.3. *Dla boole'owskich zapytań koniunkcyjych,* $q \subseteq_{set} q' \iff hom(q', q) \neq \emptyset$

Dowód opisany jest w klasycznej pracy [5]

Operacje na strukturach

Będziemy używać pewnych standardowych operacji na strukturach, opisanych między innymi w [12]. Niech A, B będą strukturami nad tą samą sygnaturą Σ. Wtedy:

- $A + B$ jest sumą rozłączną[3] A i B

[2] Multizbiór A zawiera się w B, gdy $\forall x \ A[x] \leq B[x]$

[3] to znaczy, że zastępujemy B przez izomorficzną strukturę B' taką, że $|A| \cap |B'| = \emptyset$ i bierzemy $A \cup B'$

- $A \times B$ to produkt struktur A i B, tzn. $|A \times B| = |A| \times |B|$ i dla każdej $R \in \Sigma$ (pewnej arności n) oraz $\langle a_1, ..., a_n \rangle \in A^n, \langle b_1, ..., b_n \rangle \in B^n$,

$$A \times B \models R(\langle a_1, b_1 \rangle, ..., \langle a_n, b_n \rangle) \iff A \models R(a_1, ..., a_n) \wedge B \models R(b_1, ..., b_n)$$

- Symbole $\sum, \prod$ oznaczają standardowe uogólnienie symboli $+, \times$
- Dla $t \in \mathbb{N}_+$, $tA = \sum_{i=1}^{t} A$ oraz $A^t = \prod_{i=1}^{t} A$. Ponadto, $0A$ jest strukturą pustą oraz A^0 jest jednoelementową strukturą $\{\alpha\}$ taką, że $\forall R \in \Sigma\ A^0 \models R(\alpha, ..., \alpha)$. Zauważmy, że $0A$ jest elementem neutralnym $+$ oraz A^0 jest elementem neutralnym $\times$.

Poniższy lemat jest znanym faktem dotyczącym zliczania homomorfizmów dla sum i produktów struktur, przytaczanym m. in. w [12].

Lemat 1.4. *Niech A, B, C będą strukturami i niech $t \in \mathbb{N}$. Wtedy:*

1. *Jeśli A jest spójne, to $|hom(A, B + C)| = |hom(A, B)| + |hom(A, C)|$*
2. $|hom(A, B \times C)| = |hom(A, B)| \cdot |hom(A, C)|$
3. $|hom(A, B^t)| = |hom(A, B)|^t$
4. $|hom(A + B, C)| = |hom(A, C)| \cdot |hom(B, C)|$

Graf Gaifmana

Graf Gaifmana struktury $\mathcal{D}$ to graf nieskierowany $G_{\mathcal{D}}$ zdefiniowany następująco:

- wierzchołki: $V(G_{\mathcal{D}}) = |\mathcal{D}|$
- krawędzie: $a, b \in V(G_{\mathcal{D}})$ są połączone krawędzią wtedy i tylko wtedy, gdy istnieje relacja R oraz krotka $\bar{a}$ taka, że $a, b \in \bar{a}$ oraz $\bar{a} \in R^{\mathcal{D}}$

W przypadku zapytania koniunkcyjnego oczywiście też możemy mówić o jego grafie Gaifmana - jest to graf Gaifmana jego *frozen body*.

Konwencje dotyczące algebry liniowej

W naszej pracy będziemy korzystać z oznaczeń i faktów pochodzących z algebry liniowej. Większość z nich jest standardowych, poniżej wyjaśniamy kilka nieoczywistych zapisów:

- Dla $A \subseteq \mathbb{R}^k$, $\text{span}(A)$ oznacza domknięcie liniowe A (najmniejszą przestrzeń liniową zawierającą A). Ponadto, dla zbioru $B \subseteq \mathbb{R}$, definiujemy

$$\text{span}^B(A) = \{b_1\vec{a}_1 + ... + b_n\vec{a}_n \mid n \in \mathbb{N}, b_1, ..., b_n \in B, \vec{a}_1, ..., \vec{a}_n \in A\}$$

Zatem $\text{span}(A) = \text{span}^{\mathbb{R}}(A)$.

- Dla wektorów $\vec{u}, \vec{v} \in \mathbb{R}^k$, $\langle u, v\rangle$ oznacza iloczyn skalarny $\vec{u}, \vec{v}$. Wektor $\vec{u}$ jest prostopadły do $\vec{v}$, gdy $\langle \vec{u}, \vec{v}\rangle = 0$.
- Dla wektora $\vec{u} \in \mathbb{R}^k, i \in \{1, \ldots, k\}, \vec{u}(i)$ oznacza wartość wektora $\vec{u}$ na i-tej współrzędnej.
- Dla macierzy $M \in \mathbb{R}^{k\times k}, i, j \in \{1, \ldots, k\}$, $M(i, j)$ oznacza wartość elementu w i-tym rzędzie i j-tej kolumnie M.
- Dla macierzy $M \in \mathbb{R}^{k\times k}$ i zbioru $A \subseteq \mathbb{R}^k$, $M(A) = \{M\vec{x} \mid \vec{x} \in A\}$.

Będziemy wykorzystywać parę znanych podstawowych faktów z algebry liniowej i topologii:

Fakt 1.5. *Niech $\vec{u}_1, ..., \vec{u}_n, \vec{u} \in \mathbb{Q}^k$ będą takie, że $\vec{u} \notin span\{\vec{u}_1, ..., \vec{u}_n\}$. Wtedy istnieje $\vec{z} \in \mathbb{Q}^k$ takie, że $\vec{z}$ jest prostopadły do $\vec{u}_1, ..., \vec{u}_n$ ale nie jest prostopadły do $\vec{u}$.*

Fakt 1.6. *Jeśli macierz $M \in \mathbb{R}^{k\times k}$ jest nieosobliwa, to odwzorowanie $\vec{x} \mapsto M\vec{x}$ jest homeomorfizmem (ciągłą bijekcją, której funkcja odwrotna również jest ciągła).*

Fakt 1.7. *$\mathbb{Q}^k$ jest gęstym podzbiorem $\mathbb{R}^k$, tzn., dla każdego $\vec{x} \in \mathbb{R}^k$ i $r > 0$ istnieje $\vec{y} \in \mathbb{Q}^k$ takie, że $\|\vec{x} - \vec{y}\| < r$.*

Wniosek 1.8. *Załóżmy, że $M \in \mathbb{R}^{k\times k}$ jest nieosobliwa. Wtedy istnieje $\vec{p} \in M(\mathbb{R}^k_{\geq 0}) \cap \mathbb{Q}^k$ takie, że*

$$\exists r > 0\ \forall \vec{x} \in \mathbb{R}^k \quad \|\vec{x} - \vec{p}\| < r \Rightarrow \vec{x} \in M(\mathbb{R}^k_{\geq 0}) \tag{$\star$}$$

Dowód. Z faktu 1.6 wiemy, że zbiór $M(\mathbb{R}^k_{\geq 0})$ ma niepuste wnętrze (tzn. zbiór punktów spełniających $\star$), ponieważ jest homeomorficznym obrazem zbioru o niepustym wnętrzu. Z faktu 1.7 wiemy, że ten zbiór musi zawierać punkt o współrzędnych wymiernych. $\square$

Rozdział 2

Zapytania ścieżkowe

Zawieranie zapytań ścieżkowych w semantyce zbiorowej zostało zbadane w pracy [2], gdzie wykazano, że determinacja ma dość elegancką charakteryzację w języku teorii grafów - dla pewnego grafu, w naturalny sposób reprezentującego widoki $v_1, ..., v_n$ na zapytaniu q (którego definicję przytoczymy w dalszej części pracy), zawieranie zachodzi wtedy i tylko wtedy, gdy wierzchołki reprezentujące zmienne wolne zapytania q są połączone ścieżką. W związku z czym problem ten można rozwiązać w czasie wielomanowym. Okazuje się, że ta sama charakteryzacja "działa"też w przypadku multizbiorowym!

Twierdzenie 1. *Dla zbioru zapytań ściezkowych V i zapytania ściekowego q,*

$$V \xrightarrow{set} q \iff V \xrightarrow{bag} q$$

.

To twierdzenie może być trochę zaskakujące, ponieważ inne rozważane przez nas klasy zapytań pokazują, że determinacja zachowuje się istotnie różnie w zależności od wyboru semantyki. Dowód (równoważności determinacji ze wspomnianą grafową własnością) przebiega jednak zupełnie inaczej.

Definicja 2.1. *Dla zbioru zapytań ścieżkowych V i zapytania ścieżkowego q definiujemy nieskierowany graf $G_{q,V}$ w następujący sposób:*

- *Zbiór wierzchołków $G_{q,V}$, to $\{w \in \Sigma^* \mid w \text{ jest prefiksem}^1\ q\}$. W szczególności, należy do niego słowo puste ε i całe słowo q*
- *Pomiędzy wierzchołkami w i w' istnieje krawędź, gdy $w = w'v$ lub $w' = wv$ dla pewnego $v \in V$*

Następujący fakt pochodzi z pracy [2]

[1]przypomnijmy, że zapytania ścieżkowe identyfikujemy ze słowami

Fakt 2.2. *$V \xrightarrow{set} q \iff$ istnieje ścieżka z ε do q w $G_{q,V}$*

Żeby udowodnić twierdzenie 1, wystarczy pokazać następującą równoważność:

Lemat 2.3. *$V \xrightarrow{bag} q \iff$ istnieje ścieżka z ε do q w $G_{q,V}$*

Resztę tego rozdziału poświęcimy dowodzeniu lematu 2.3.

2.1 Dowód implikacji $\Leftarrow$

Załóżmy, że q i V są takie, że w $G_{q,V}$ istnieje ścieżka z ε do q. Niech m będzie długością takiej ścieżki. Zatem istnieje:

- ciąg $w_0, w_1, ..., w_m$ będących prefiksami q, przy czym $w_0 = \varepsilon, w_m = q$
- ciąg liczb $\epsilon_1, ..., \epsilon_m \in \{-1, 1\}$
- ciąg $v_{p_1}, ..., v_{p_m} \in V$

takie, że dla każdego $j \in \{1, \dots, m\}$ jeden z poniższych warunków jest spełniony:

a) $\epsilon_j = 1 \wedge w_j = w_{j-1} v_{p_j}$

b) $\epsilon_j = -1 \wedge w_j v_{p_j} = w_{j-1}$

Pokażemy, że $q(D)$ można przedstawić jako wynik pewnej funkcji f (niezależnej od wyboru D) o argumentach $v_1(D), v_2(D), ..., v_k(D)$. Stąd już wynika, że $V \xrightarrow{\text{bag}} q$ (obserwacja 1.1).

2.1.1 q-spacery i ich redukcje

Niech $\bar{\Sigma} = \Sigma \cup \Sigma^{-1}$ będzie naszym nowym alfabetem, przy czym $\Sigma^{-1} = \{R^{-1} \mid R \in \Sigma\}$.

Definicja 2.4. *Niech $w \in \bar{\Sigma}^*$ i $w = A_1^{\iota_1} ... A_k^{\iota_k}$ dla pewnych $A_1, ..., A_k \in \Sigma$ oraz $\iota_1, ..., \iota_k \in \{1, -1\}$. Wtedy w nazywamy q-*spacerem *gdy:*

(1) dla każdego $i \in \{1, \dots, k\}$ zachodzi $0 \leq \sum_{j=1}^{i} \iota_j \leq |q|$;

(2) $\sum_{j=1}^{k} \iota_j = |q|$;

(3) dla każdego $i \in \{1, \dots, k\}$ zachodzi $A_i = \begin{cases} Q_{s_i+1} & gdy\ \iota_i = 1 \\ Q_{s_i} & gdy\ \iota_i = -1 \end{cases}$

gdzie $s_i = \sum_{j=1}^{i-1} \iota_j$ oraz Q_j jest j-tym symbolem q.

Mniej formalnie, q-*spacer* jest zapisem pewnego spaceru na słowie q, który zaczyna się na początku tego słowa (przed pierwszą literą) kończy na jego końcu (za ostatnią literą), a każdy krok polega na "przeskoczeniu" jednej litery w prawo lub w lewo. Każdy taki krok powoduje zapis "przeskoczonej" litery, przy czym gdy skaczemy w lewo, zapisujemy ją z "-1" w górnym indeksie.

Nasza ścieżka w grafie $G_{q,v}$ w naturalny sposób generuje q-*spacer* jako $(v_{p_1})^{\epsilon_1}...(v_{p_m})^{\epsilon_m}$ (przy czym słowo w^{-1} rozumiemy jako słowo powstałe poprzez odwrócenie słowa w i dopisania $^{-1}$ do wszystkich liter). Dla jasności zilustrujemy to następującym przykładem:

Przykład 2.5. *Załóżmy, że* $q = ABCD$ *oraz* $V = \{ABC, BC, BCD\}$. *Wtedy w* $G_{q,V}$ *istnieje ściezka* $\varepsilon \to ABC \to A \to ABCD$, *która generuje* q-spacer $(ABC)(BC)^{-1}(BCD) = ABCC^{-1}B^{-1}BCD$

Pokażemy, że q-*spacer* możemy uprościć do słowa q za pomocą następujących redukcji:

Definicja 2.6. *Relacje* $\xrightarrow[+/-]{}$, $\xrightarrow[-/+]{}$ *definiujemy jako najmniejsze relacje, które dla każdego* $w, w' \in \bar{\Sigma}$, $A \in \Sigma$, *spełniają:*

(1) $wAA^{-1}w' \xrightarrow[+/-]{} ww'$

(2) $wA^{-1}Aw' \xrightarrow[-/+]{} ww'$

Definiujemy również relacje $\xrightarrow[+/-]{*}$, $\xrightarrow[-/+]{*}$ *jako przechodnie, zwrotne domkniecia relacji* $\xrightarrow[+/-]{}$ *i* $\xrightarrow[-/+]{}$.

Lemat 2.7. *Jeśli* $w \in \bar{\Sigma}^*$ *jest* q-spacerem, *to* $w \xrightarrow[+/-]{*} q$ *oraz* $w \xrightarrow[-/+]{*} q$.

Dowód. Załóżmy, że $w \in \bar{\Sigma}$ jest q-spacerem. Udowodnimy, że $w \xrightarrow[+/-]{*} q$. Dowód, że $w \xrightarrow[-/+]{*} q$ jest analogiczny. Dowód jest indukcyjny względem długości w. Ponieważ w jest q-spacerem, $|w| \geq q$. Rozważmy dwa przypadki:

1. $|w| = |q|$. Wtedy $w = q$, więc jest OK

2. $|w| > |q|$. Wtedy istnieje i takie, że $\iota_i = -1$ (ι_i takie, jak w definicji 2.4). Ale z definicji 2.4(1) wiemy, że $\iota_1 = 1$, więc istnieje $j < i$ takie, że $\iota_j = 1, \iota_{j+1} = -1$. Zatem z punktu (3) tej samej definicji, otrzymujemy, że dla pewnego $A \in \Sigma$, $w_j = A$ oraz $w_{j+1} = A^{-1}$. Zatem w jest postaci $uAA^{-1}u'$ dla pewnych słów u, u'. Łatwo sprawdzić, że uu' również jest q-spacerem, więc z założenia indukcyjnego, $uu' \xrightarrow[+/-]{*} q$. Oczywiście $w \xrightarrow[+/-]{} uu'$, zatem również $w \xrightarrow[+/-]{*} q$.

$\square$

2.1.2 Bazy danych a przekształcenia liniowe

Definicja 2.8. *Niech $R \in \Sigma$. Niech D będzie bazą danych o uniwersum $\{a_1, ..., a_n\}$. Wtedy M_R^D jest macierzą incydencji relacji R w strukturze D, tzn. $M_R^D \in \mathbb{R}^{n \times n}$ oraz*

$$M_R^D(i,j) = \begin{cases} 1 & gdy\ D \models R(a_i, a_j) \\ 0 & w\ p.\ p. \end{cases}$$

Definicja 2.9. *Niech $w \in \Sigma^*$. Wtedy określamy $M_w^D \in \mathbb{R}^{n \times n}$ w następujący induktywny sposób:*

1. *M_ε^D jest macierzą identycznościową wymiaru $n \times n$.*
2. *Dla $R \in \Sigma, w \in \Sigma^*$, $M_{Rw}^D = M_R^D M_w^D$.*

Zauważmy, że dla $a_i, a_j \in D, w \in \Sigma^*$, $w(D)[a_i, a_j]$ jest liczbą ścieżek z a_i do a_j przechodzących przez krawędzie o etykietach kolejno: $w_1, w_2, ..., w_{|w|}$. Dobrze znanym faktem z teorii grafów jest, że:

Fakt 2.10. *Dla $w \in \Sigma^*$, $w(D)[a_i, a_j] = M_w^D(i,j)$.*

Zatem macierze $M_{v_1}^D, ..., M_{v_n}^D$ są jednoznacznie wyznaczone przez $v_1(D), ..., v_n(D)$. Ponadto, M_q^D jednoznacznie wyznacza $q(D)$. Zatem, by dokończyć dowód, wystarczy pokazać, że M_q^D można wyliczyć jako funkcję o argumentach $M_{v_1}^D, ..., M_{v_n}^D$.

Uwaga 2.11. *Załóżmy, że baza danych D jest taka, że wszystkie macierze $\{M_R^D \mid R \in \Sigma\}$ są odwracalne. Łatwo wtedy sprawdzić, że $M_q^D = (M_{v_{p_1}}^D)^{\epsilon_1} ... (M_{v_{pm}}^D)^{\epsilon_m}$, zatem M_q^D przedstawia się jako funkcja $M_{v_1}^D, ..., M_{v_n}^D$. Niestety, oczywiście, nie wszystkie macierze są odwracalne. W ogólnym przypadku, potrzebne będą bardziej wyrafinowane metody.*

Jak wiemy z algebry liniowej, macierze możemy identyfikować z przekształceniami liniowymi, czyli pewnymi funkcjami $\mathbb{R}^n \to \mathbb{R}^n$. Z kolei na funkcje możemy patrzeć jako na binarne relacje. Choć nie wszystkie przekształcenia liniowe są odwracalne (i, w związku z tym, również nie wszystkie macierze), to relacje zawsze możemy odwracać! Okaże się to kluczowe w dalszej części dowodu.

Definicja 2.12.

1. *Przez I będziemy oznaczać relację identycznościową: $I = \{(x, x) \mid x \in \mathbb{R}^n\}$.*
2. *Dla macierzy $M \in \mathbb{R}^{n \times n}$ definiujemy przekształcenie liniowe $h_M : \mathbb{R}^n \to \mathbb{R}^n$ jako $h_M(v) = Mv$*
3. *Dla funkcji f definiujemy $\overline{f}$ jako relację równą wykresowi f, to znaczy $\overline{f} = \{(x, y) \mid f(x) = y\}$[2].*

[2] Choć z punktu widzenia teoriomnogościowego, f i $\overline{f}$ wydają się być tym samym, wprowadzamy powyższe oznaczenie, by jasne było, jak określone są operacje złożenia i odwrotności - różnią się w zależności od tego, czy mówimy o funkcji czy o relacji!

4. *Dla* $R \in \Sigma$ *definiujemy* $H_R = \overline{h_{M_R^D}}$ *oraz* $H_{R^{-1}} = H_R^{-1}$
5. *Dla* $w \in \bar{\Sigma}^*$ *definiujemy* H_w *indukcyjnie:*
   - $H_\varepsilon = I$
   - $H_{\alpha w} = H_w H_\alpha$ *dla* $\alpha \in \bar{\Sigma}, w \in \bar{\Sigma}^*$

Obserwacja 2.13. *Dla* $w \in \bar{\Sigma}^*$, $H_w = \overline{h_{M_w^D}}$ *oraz* $H_{w^{-1}} = (H_w)^{-1}$

Dowód. Sprawdzamy, że dla dowolnych $w, w' \in \bar{\Sigma}^*$ zachodzi $H_{ww'} = H_{w'} H_w = \overline{h_{M_{w'}^D}}\,\overline{h_{M_w^D}} = \overline{h_{M_w^D} \circ h_{M_{w'}^D}} = \overline{h_{M_{ww'}^D}}$. Ponadto $H_\varepsilon = I = \overline{h_{M_\varepsilon^D}}$. Reszta dowodu to prosta indukcja. $\square$

Znanym faktem jest, że przyporządkowanie $M \mapsto h_M$ jest *1-1*. Także przyporządkowanie $f \mapsto \bar{f}$ jest *1-1*. Zatem, by wykazać, że M_q^D przedstawia się jako funkcja argumentów $M_{v_1}, ..., M_{v_n}$, wystarczy pokazać, że H_q przedstawia się jako funkcja argumentów $H_{v_1}, ..., H_{v_n}$.

2.1.3 Wykorzystanie dwóch redukcji q-spacerów

Zaczniemy od bardzo prostego lematu:

Lemat 2.14. *Niech* $f : \mathbb{R}^n \to \mathbb{R}^n$. *Wtedy* $\bar{f}\ (\bar{f})^{-1} \supseteq I$ *oraz* $(\bar{f})^{-1}\ \bar{f} \subseteq I$.

Dowód. $\bar{f}\ (\bar{f})^{-1} = \{(x, y) \mid \exists z\ f(x) = z \wedge f(y) = z\} = \{(x, y) \mid f(x) = f(y)\} \supseteq I$

$$(\bar{f})^{-1}\ \bar{f} = \{(x, y) \mid \exists z\ f(z) = x \wedge f(z) = y\} = \{(x, x) \mid \exists z\ f(z) = x\} \subseteq I$$ $\square$

Nadszedł czas, by połączyć wątki z ostatnich dwóch sekcji:

Lemat 2.15. *Dla* $u, u' \in \bar{\Sigma}^*$, *zachodzi:*

1. *jeśli* $u \xrightarrow[+/-]{} u'$, *to* $H_u \subseteq H_{u'}$;
2. *jeśli* $u \xrightarrow[-/+]{} u'$, *to* $H_u \supseteq H_{u'}$.

Dowód. 1. Jeśli $u \xrightarrow[+/-]{} u'$, to istnieją $w, w' \in \bar{\Sigma}^*$ oraz $R \in \Sigma$ takie, że $u = wRR^{-1}w'$ oraz $u' = ww'$. Wtedy, używając lematu 2.15:

$$H_u = H_{wRR^{-1}w'} = H_{w'} H_R^{-1} H_R H_w = H_{w'} \overline{h_{M_R}}^{-1} \overline{h_{M_R}} H_w \subseteq H_{w'} I H_w = H_{ww'} = H_{u'}$$

2. Jeśli $u \xrightarrow[-/+]{} u'$, to istnieją $w, w' \in \bar{\Sigma}^*$ oraz $R \in \Sigma$ takie, że $u = wR^{-1}Rw'$ oraz $u' = ww'$. Wtedy, ponownie z lematu 2.15:

$$H_u = H_{wR^{-1}Rw'} = H_{w'} H_R H_R^{-1} H_w = H_{w'} \overline{h_{M_R}}\, \overline{h_{M_R}}^{-1} W H_w \supseteq H_{w'} I H_w =$$

$$= H_{ww'} = H_{u'}$$

$\square$

Dzięki lematom 2.7 i 2.15 otrzymujemy dwie aproksymacje H_q:

Wniosek 2.16. *Jeśli w jest q-spacerem, to $H_q \subseteq H_w \subseteq H_q$, zatem $H_w = H_q$.*

Przypomnijmy, że $(v_{p_1})^{\epsilon_1}(v_{p_2})^{\epsilon_2}\dots(v_{p_m})^{\epsilon_m}$ jest q-spacerem. Zatem, z powyższego wniosku wynika $H_q = H^{\epsilon_m}_{v_{p_m}}\dots H^{\epsilon_2}_{v_{p_2}}H^{\epsilon_1}_{v_{p_1}}$, co kończy dowód implikacji $\Leftarrow$ lematu 2.3.

2.2 Dowód implikacji $\Rightarrow$

Ten dowód także pochodzi z pracy [2], gdzie dowodzony był analogiczny lemat w semantyce zbiorowej. Okazuje się, że ta sama konstrukcja działa również w przypadku multizbiorowym.

Załóżmy, że nie istnieje ścieżka z ε do q w $G_{q,V}$. Pokażemy, że w takim razie V nie determinuje q. Zdefiniujmy strukturę D w następujący sposób:

- $|D| = \{[w,j] \mid w \text{ jest prefiksem } q, j \in \{0,1\}\}$
- Dla $[w,i],[u,j] \in D, R \in \Sigma$ zachodzi $D \models R([w,i],[u,j])$ wtedy i tylko wtedy, gdy $u = wR$ oraz $i = j$.

Zatem D jest izomorficzne z $q + q$, czyli sumą dwóch *frozen body* q. Łatwo sprawdzić z definicji, że $([\varepsilon,0],[q,0]) \in q(D)$ z krotnością 1.

Obserwacja 2.17. *Niech $z \in \Sigma^*$ będzie dowolnym słowem (z którym utożsamiamy również zapytanie ścieżkowe), niech $w,u \in G_{q,V}$ oraz $i,j \in \{0,1\}$. Wtedy $([w,i],[u,j]) \in z(D)$ wtedy i tylko wtedy, gdy:*

$$i = j \land u = wz$$

Ponadto inkluzja zawsze jest z krotnością 1,

Dowód. Prosta indukcja. □

Zdefiniujemy relację równoważności $\sim$ na $G_{q,V}$: $w \sim v$ wtedy i tylko wtedy, gdy zachodzi jeden z dwóch warunków:

- zarówno w jak i v są osiągalne z ε w grafie $G_{q,V}$
- zarówno w jak i v nie są osiągalne z ε w grafie $G_{q,V}$

Oczywiście jest to relacja równoważności z dwiema klasami, ponadto $\varepsilon \not\sim q$.

Określimy też drugą strukturę D', taką że $|D'| = |D|$ oraz, dla dowolnych $u,w \in G_{q,V}$, $R \in \Sigma$, $i,j \in \{0,1\}$:

- Jeśli $i = j$, to $D' \models R([w, i], [u, j])$ wtedy i tylko wtedy, gdy $u = wR$ oraz $u \sim w$
- Jeśli $i \neq j$, to $D' \models R([w, i], [u, j])$ wtedy i tylko wtedy, gdy $u = wR$ oraz $u \not\sim w$

Uwaga 2.18. *D' również jest izomorficzne z $q+q$, choć nie jest to ten sam izomorfizm, co w przypadku D.*

Obserwacja 2.19. *Niech $z \in \Sigma^*$ będzie dowolnym słowem (z którym utożsamiamy również zapytanie ścieżkowe), niech $w, u \in G_{q,V}$ oraz $i, j \in \{0, 1\}$. Wtedy $([w, i], [u, j]) \in z(D')$ wtedy i tylko wtedy, gdy:*

$$(i = j \iff w \sim u) \wedge u = wz$$

Ponadto należenie zawsze jest z krotnością 1,

Dowód. Ponownie prosta indukcja. □

W szczególności, ponieważ $\varepsilon \not\sim q$, $([\varepsilon, 0], [q, 0]) \notin q(D')$. Zatem $q(D) \neq q(D')$

Przypomnijmy, że jeśli dla pewnych $u, w \in G_{q,V}$ oraz pewnego $v \in V$ zachodzi $u = wv$, to u, w są połączone krawędzią i w konsekwencji $u \sim w$. Z tego oraz z obserwacji 2.17 i 2.19 wynika, że dla każdego $v \in V$, $v(D) = v(D')$. Skonstruowaliśmy zatem kontrprzykład przeczący, że $V \xrightarrow{\text{bag}} q$.

Rozdział 3

Unie zapytań koniunkcyjnych (UCQ)

W przypadku zapytań ścieżkowych udowodniliśmy, że determinacja w semantyce zbiorowej jest równoważna determinacji multizbiorowej. Przypadek UCQ pokazuje jednak, że nie jest to regułą. Co więcej, jesteśmy w stanie podać przykład, dla którego zachodzi determinacja multizbiorowa, a zbiorowa już nie!

Przykład 3.1. *Niech $q = \exists x\ R(x)$ oraz niech V składa się z następujących zapytań:*

$$v_1 = \exists x\ P(x) \qquad\qquad v_2 = \exists x\ P(x) \lor \exists x\ R(x)$$

Wtedy $V \overset{set}{\not\longrightarrow} q$. Kontrprzykład to dwie jednoelementowe bazy danych D, D':

- $D \models \exists x\ P(x)$, $D \models \neg\exists x\ R(x)$
- $D' \models \exists x\ P(x)$, $D' \models \exists x\ R(x)$

Tutaj (w sensie zbiorowym) $v_1(D) = v_1(D'), v_2(D) = v_2(D')$, ale $q(D) \neq q(D')$.

Z drugiej strony, $V \xrightarrow{bag} q$, ponieważ dla każdej bazy danych D zachodzi $q(D) = v_2(D) - v_1(D)$.

Przejdźmy do naszego głównego twierdzenia dotyczącego UCQ:

Twierdzenie 2. *Problem, czy dla danego V będącego skończonym zbiorem boole'owskich UCQ oraz dla jeszcze jednego boole'owskiego UCQ q, zachodzi $V \xrightarrow{bag} q$ jest nierozstrzygalny.*

Warto zaznaczyć, że powyższe twierdzenie nie jest prawdą w przypadku zbiorowym - nawet dla unarnych UCQ problem determinacji UCQ jest rozstrzygalny!

Przejdźmy do dowodu twierdzenia 2:

Dowód. Przypomnijmy (powszechnie znaną) definicję 10. Problemu Hilberta:

Problem 3.2. *Dla danego wielomianu o współczynnikach całkowitych (być może wielomianu wielu zmiennych) rozstrzygnąć, czy posiada on miejsce zerowe o wartościach naturalnych.*

Wiadomo, że ten problem jest nierozstrzygalny.

Instancje tego problemu możemy rozumieć jako zbiór jednomianów o współczynnikach całkowitych. Dla jednomianu m przez $c(m)$ będziemy oznaczać jego całkowitoliczbowy współczynnik, a dodatkowo dla zmiennej x, przez $m(x)$ oznaczamy stopień, w którym x występuje w m (jeśli x nie występuje w m, to $m(x) = 0$).

Nasz dowód będzie się opierał na redukcji z dopełnienia 10. Problemu Hilberta. Weźmy instancję tego problemu: $I = \{m_1, m_2, ..., m_k\}$. Pokażemy algorytm, który wygeneruje sygnaturę Σ oraz boole'owskie UCQ q i zbiór boole'owskich UCQ V taki, że $V \xrightarrow{\text{bag}} q$ wtedy i tylko wtedy, gdy I **nie ma** rozwiązania w zbiorze liczb naturalnych.

Niech $x_1, x_2, ..., x_n$ oznaczają wszystkie zmienne jednomianów $m_1, ..., m_k$. Tworzymy sygnaturę Σ następująco: składa się z nullarnych[1] predykatów H, C oraz unarnych predykatów $X_1(x), ..., X_n(x)$. Dla struktury D i symbolu $R \in \Sigma$ definiujemy D_R jako liczbę atomów relacji R występujacych w D. Zwróćmy uwagę że ponieważ H, C są nullarne, D_H oraz $D_C \in \{0, 1\}$.

Naszym celem jest skonstruować takie V, które zapewni, że dla dowolnych dwóch baz D, D' spełniających $V(D) = V(D')$, zachodzi $D_{X_i} = D'_{X_i}$ dla każdego i. Zatem, jeśli D, D' miałyby być kontrprzykładem determinacji, muszą się czymś różnić i jedyny sposób, by to osiągnąć jest taki, by różniły się na predykatach H, C. Do V będzie też należało zapytanie, które w takim wypadku wymusi równość pewnych dwóch wielomianów w punkcie $(D_{X_1}, ..., D_{X_n})$, co będzie implikowało, że wielomian podany na wejściu ma w tym punkcie miejsce zerowe.

Na początek dla danego jednomianu m zdefiniujmy następujące zapytanie koniunkcyjne:

$$\Phi_m = \exists^* \bigwedge_{X_i \in \Sigma} \bigwedge_{j=1}^{m(x_i)} X_i(y_{i,j})$$

Gdzie kwantyfikator $\exists^*$ wiąże wszystkie zmienne wolne występujące w podanej formule.

Definicja 3.3. *Niech $m \in I$. Wtedy:*

- *Dla krotki $\bar{a} = \langle a_1, ..., a_n \rangle \in \mathbb{N}^n$ definiujemy $m_{\bar{a}}$ jako wartość m po podstawieniu $x_i := a_i$ dla każdego $i \in \{1, \ldots, n\}$.*

[1] tzn. zero-argumentowych

- *Dla struktury D, $m_D = m_{\langle D_{X_1},\ldots,D_{X_n}\rangle}$*

Lemat 3.4. *Dla dowolnej bazy D oraz $m \in I$:*

$$m_D = c(m) \cdot \Phi_m(D)$$

Dowód wynika bezpośrednio z lematu 1.4.

Niech $P = \{m \in I \mid c(m) > 0\}$, $N = \{m \in I \mid c(m) < 0\}$. Zdefiniujemy następujące UCQ:

$$\Psi_P = \bigvee_{m\in P} \bigvee_{i=1}^{c(m)} \Phi_m \wedge H, \qquad \Psi_N = \bigvee_{m\in N} \bigvee_{i=1}^{-c(m)} \Phi_m \wedge C.$$

Lemat 3.5. *Dla dowolnej bazy D zachodzi:*

$$D_H \cdot \sum_{m\in P} m_D = \Psi_P(D).$$

Dowód.

$$\begin{aligned}
\Psi_P(D) &= \sum_{m\in P} \sum_{i=1}^{c(m)} (\Phi_m \wedge H)(D) && \text{(lemat 1.4)}\\
&= \sum_{m\in P} \sum_{i=1}^{c(m)} D_H \cdot \Phi_m(D)\\
&= D_H \cdot \sum_{m\in P} \sum_{i=1}^{c(m)} \Phi_m(D)\\
&= D_H \cdot \sum_{m\in P} c(m) \cdot \Phi_m(D)\\
&= D_H \cdot \sum_{m\in P} m_D && \text{(lemat 3.4)}
\end{aligned}$$

□

Lemat 3.6. *Dla dowolnej bazy D zachodzi:*

$$D_C \cdot \sum_{m\in N} m_D = -\Psi_N(D)$$

Dowód. Analogiczny do dowodu lematu 3.5. □

Przejdźmy do definicji q oraz V. Niech $q = H$ oraz niech V składa się z następujących zapytań:

- $v_1 = H \vee C$
- dla każdego $i \in \{1, \ldots, n\}$: $v_{X_i} = \exists y X_i(y)$

- $v_I = \Psi_P \vee \Psi_N$

Powyższa definicja V zapewnia następujące właściwości:

Lemat 3.7. *Dla dowolnych D, D' spełniających $V(D) = V(D')$ oraz $q(D) \neq q(D')$ zachodzi:*

$$D_{X_i} = D'_{X_i}, \quad D_H = D'_C, \quad D_C = D'_H, \quad D_H \neq D_C$$

Dowód. $D_{X_i} = v_{X_i}(D) = v_{X_i}(D') = D'_{X_i}$. Ponadto z równości $v_1(D) = v_1(D')$ otrzymujemy następujące możliwości:

1. $D_H = D'_C, \quad D_C = D'_H, \quad D_H \neq D_C$ $\qquad (V_1(D) = 1)$

2. $D_H = D'_H, \quad D_C = D'_C, \quad D_H \neq D_C$ $\qquad (V_1(D) = 1)$

3. $D_H = D'_H = D_C = D'_C = 0$ $\qquad (V_1(D) = 0)$

4. $D_H = D'_H = D_C = D'_C = 1$ $\qquad (V_1(D) = 2)$

Ponieważ jednak $q(D) = D_H$, $q(D') = D'_H$, tylko pierwsza możliwość spełnia warunek $q(D) \neq q(D')$. □

By dokończyć dowód twierdzenia 2 pozostaje nam teraz pokazać, że:

Lemat 3.8. *Istnieje para struktur D, D' spełniająca $V(D) = V(D'), q(D) \neq q(D')$ wtedy i tylko wtedy, gdy I ma pierwiastek w zbiorze liczb naturalnych.*

Dowód. ($\Leftarrow$) Niech $\bar{a} \in \mathbb{N}^n$ będzie pierwiastkiem I. Wtedy niech D, D' będą takie, że:

- $D_H = 1$, $D'_H = 0$, $D_C = 0$, $D'_C = 1$
- $D_{X_i} = D'_{X_i} = a_i$

Z lematów 3.5, 3.6 otrzymujemy, że

$$v_I(D) - v_I(D') = \sum_{m \in P} m_{\bar{a}} + \sum_{m \in N} m_{\bar{a}} = \sum_{m \in I} m_{\bar{a}} = 0$$

Zatem $V(D) = V(D')$.

($\Rightarrow$) Weźmy D, D' takie, że $V(D) = V(D')$ oraz $q(D) \neq q(D')$. Wtedy $D_H = D'_C, D_C = D'_H, D_H \neq D_C$. Bez straty ogólności, $D_H = D'_C = 1$, $D'_H = D_C = 0$. Pokażemy, że $\sum_{m \in I} m_D = 0$.

Z lematów 3.5, 3.6 otrzymujemy:

$$
\begin{aligned}
V_I(D) &= V_I(D') \\
\Psi_P(D) + \Psi_N(D) &= \Psi_P(D') + \Psi_N(D') \\
\Psi_P(D) &= \Psi_N(D') \\
\sum_{m \in P} m_D &= -\sum_{m \in N} m_D \\
\sum_{m \in P} m_D + \sum_{m \in N} m_D &= 0 \\
\sum_{m \in I} m_D &= 0
\end{aligned}
$$

□

□

Rozdział 4

Boole'owskie zapytania koniunkcyjne

W poprzednim rozdziale pokazaliśmy nierozstrzygalność determinacji nawet dla boole'owskich UCQ. Co jednak, gdybyśmy dodatkowo zażądali, by V oraz q składały się z zapytań koniunkcyjnych? Okazuje się, że wtedy problem staje się rozstrzygalny, o „niezbyt wysokiej" złożoności `P-SPACE` (dokładniej: Σ_2^P). Przejdźmy do omówienia najważniejszego wyniku w tej pracy:

Twierdzenie 3. *Niech $V_0 = \{v_1, ..., v_n\}$ będzie zbiorem boole'owskich zapytań koniunkcyjnych i niech q będzie boole'owskim zapytaniem koniunkcyjnym. Wtedy problem, czy $V_0 \xrightarrow{bag} q$ jest rozstrzygalny.*

Pierwsze trzy sekcje tego rozdziału poświęcimy dowodowi twierdzenia 3. Od teraz ustalamy V_0 oraz q jako instancję problemu opisanego w twierdzeniu. Naszym celem będzie pokazanie warunku równoważnego $V_0 \xrightarrow{\text{bag}} q$, który jest w oczywisty sposób rozstrzygalny.

4.1 Warunek równoważny

Definicja 4.1. *Niech $V = \{v \in V_0 \mid q \subseteq_{set} v\}$. Ponadto niech $V_q = V \cup \{q\}$.*

V jest podzbiorem V_0 składającym się z zapytań, które faktycznie będą nas interesować. Istotnie, pozostałe zapytania w ogóle nie wpływają na determinację:

Lemat 4.2. *Niech V, V_0, q będą takie, jak powyżej. Wtedy $V \xrightarrow{bag} q \iff V_0 \xrightarrow{bag} q$.*

Dowód.

1. $\Rightarrow$. Oczywiste, bo $V \subseteq V_0$.

2. $\Leftarrow$. Załóżmy, że $V \not\xrightarrow{\text{bag}} q$. Pokażemy, że również $V_0 \not\xrightarrow{\text{bag}} q$. Niech D, D' będą kontrprzykładem na $V \xrightarrow{\text{bag}} q$, to znaczy D, D' spełniają:

   (a) $\forall v \in V \ v(D) = v(D')$

   (b) $q(D) \neq q(D')$

   Rozważmy struktury $q \times D, q \times D'$. Z lematu 1.4, spełniają one:

   (a) $\forall v \in V \ v(q \times D) = v(q)v(D) = v(q)v(D') = v(q \times D')$

   (b) $\forall v \in V_0 \setminus V \ v(q \times D) = v(q)v(D) = 0 = v(q)v(D') = v(q \times D')$. Wynika to z definicji V oraz faktu, że $q \subseteq_{\text{set}} v \iff v(q) \neq 0$.

   (c) $q(q \times D) = q(q) \times q(D) \neq q(q) \times q(D') = q(q \times D')$. Tutaj należy zaznaczyć, że korzystamy z faktu, iż $q(q) > 0$ (ponieważ $q(q) = |hom(q, q)|$ oraz oczywiście $id \in hom(q, q)$)

   Zatem $q \times D, q \times D'$ stanowią przykład dowodzący, że $V_0 \not\xrightarrow{\text{bag}} q$.

□

Od teraz będziemy już tylko badać determinację $V \xrightarrow{\text{bag}} q$.

Obserwacja 4.3. *Dla dowolnej bazy D oraz $v \in V$, jeśli $v(D) = 0$, to również $q(D) = 0$.*

Definicja 4.4. *Niech $W = \{w_1, ..., w_k\}$ będzie zbiorem[1] wszystkich spójnych składowych[2] zapytań ze zbioru V_q. Innymi słowy, W jest minimalnym zbiorem struktur takim, że dla każdego $v \in V_q$ i dla każdej spójnej składowej v' należącej do v istnieje $w \in W$ izomorficzne z v'. Od teraz k będzie zawsze oznaczało moc zbioru W.*

Zapytania ze zbioru W będą dla nas *zapytaniami bazowymi*. Bazowymi, także w rozumieniu z algebry liniowej:

Obserwacja 4.5. *Niech $v \in V_q$. Wtedy $v = \sum_{i=1}^{k} a_i w_i$ dla pewnych $a_1, ..., a_k \in \mathbb{N}$.*

Definicja 4.6. *Dla zapytania $v \in V_q$ definiujemy jego reprezentację wektorową jako $\vec{v} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$, gdzie $a_1, ..., a_k$ są takie jak w obserwacji 4.5.*

Zatem zapytania będziemy postrzegać jako wektory w k-wymiarowej przestrzeni liniowej.

Obserwacja 4.7. *Dla struktury D i zapytania $v \in V_q$ zachodzi $v(D) = \prod_{i=1}^{k} w_i(D)^{\vec{v}(i)}$.*

[1] Jest to zbiór, zatem każda spójna składowa występuje w nim co najwyżej raz

[2] ponieważ zapytania utożsamiamy z grafami, możemy dla nich definiować spójne składowe zgodnie z tym utożsamieniem

Dowód wynika z lematu 1.4.

Możemy teraz postawić nasz główny lemat:

Lemat 4.8. *$V \xrightarrow{bag} q$ wtedy i tylko wtedy, gdy $\vec{q} \in span\{\vec{v} \mid v \in V\}$*

Oczywiście, jeśli udowodnimy ten lemat, będzie to koniec dowodu twierdzenia, ponieważ wszystkie kroki potrzebne do policzenia $\vec{q}$, $\{\vec{v} \mid v \in V\}$ oraz sprawdzenie, czy $\vec{q} \in \mathrm{span}\{\vec{v} \mid v \in V\}$, są rozstrzygalne.

Zanim przejdziemy do dowodu lematu 4.8 zobaczmy następujący przykład:

Przykład 4.9. *Niech w_1, w_2, w_3 będą pewnymi niepustymi, parami nieizomorficznymi strukturami oraz niech:*

$$q = w_1 + w_2 + 2w_3$$
$$v_1 = 2w_1 + w_2 + 3w_3$$
$$v_2 = 5w_1 + 2w_2 + 7w_3$$

Wtedy dla dowolnej bazy D:

$$q(D) = w_1(D)w_2(D)w_3(D)^2$$
$$v_1(D) = w_1(D)^2w_2(D)w_3(D)^3$$
$$v_2(D) = w_1(D)^5w_2(D)^2w_3(D)^7$$

O ile $v_2(D) \neq 0$, to $q(D) = v_1(D)^3/v_2(D)$, więc q jest jednoznacznie zdeterminowane przez v_1, v_2. Ta równość odpowiada równości reprezentacji wektorowych: $\vec{q} = 3\vec{v}_1 - \vec{v}_2$. Z kolei, jeśli $v_2(D) = 0$, to pewna z wartości $w_1(D), w_2(D), w_3(D)$ jest równa 0, zatem $q(D) = 0$, czyli znów q jest jednoznaczenie wyznaczone.

4.2 Dowód lematu 4.8 (⇐)

Niech D, D' będą strukturami spełniającymi $v(D) = v(D')$ dla każdego $v \in V_0$. Chcemy pokazać, że $q(D) = q(D')$. Rozważymy dwa przypadki:

1. $\exists v \in V \ v(D) = 0$.

   Wtedy oczywiście również $v(D') = 0$. Z obserwacji 4.3 wynika, że $q(D) = 0$. Analogicznie $q(D') = 0$, zatem $q(D) = q(D')$.

2. $\forall v \in V \ v(D) \neq 0$.

Weźmy $\alpha_1, ..., \alpha_k \in \mathbb{R}$ takie, że $\vec{q} = \sum_{i=1}^{|V|} \alpha \vec{v}_i$.

$$\begin{aligned} q(D) &= \prod_{i=1}^{k} w_i(D)^{\vec{q}(i)} && \text{(z obserwacji 4.7)} \\ &= \prod_{i=1}^{k} w_i(D)^{\sum_{j=1}^{|V|} \alpha_j \vec{v}_j(i)} \\ &= \prod_{i=1}^{k} \prod_{j=1}^{|V|} \left(w_i(D)^{\vec{v}_j(i)} \right)^{\alpha_j} \\ &= \prod_{j=1}^{|V|} \left(\prod_{i=1}^{k} w_i(D)^{\vec{v}_j(i)} \right)^{\alpha_j} \\ &= \prod_{j=1}^{|V|} v_j(D)^{\alpha_j} && \text{(znów z obserwacji 4.7)} \end{aligned}$$

Ponieważ dla $j \in \{1, ..., k\}$ mamy $v_j(D) > 0$, wyrażenie $\prod_{j=1}^{|V|} v_j(D)^{\alpha_j}$ jest poprawne, nawet jeśli pewne α_j nie są liczbami naturalnymi.

Analogiczne pokazujemy, że $q(D') = \prod_{j=1}^{|V|} v_j(D')^{\alpha_j}$. Wiemy, że dla $j \in \{1, ..., k\}$ zachodzi $v_j(D) = v_j(D')$, zatem $q(D') = \prod_{j=1}^{|V|} v_j(D)^{\alpha_j} = q(D)$.

4.3 Dowód lematu 4.8 ($\Rightarrow$)

W tej sekcji zakładamy, że $\vec{q} \notin \text{span}\{\vec{v} \mid v \in V\}$. Pokażemy, że wtedy $V_0 \not\xrightarrow{\text{bag}} q$. W tym celu znajdzemy parę struktur D, D' będącą kontrprzykładem determinacji, czyli spełniającej:

(A) $q(D) \neq q(D')$

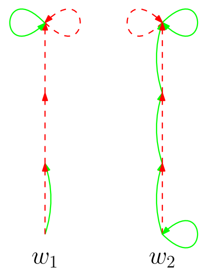
(B) $\forall v \in V \;\; v(D) = v(D')$

4.3.1 Podstawowe narzędzia

Na pierwszy rzut oka trudno jest wskazać, jak szukać pary struktur spełniającej powyższe warunki. Kluczowym pomysłem okazało się, by ograniczyć poszukiwania do pewnego podzbioru struktur utożsamianego ze stożkiem w k-wymiarowej[3] przestrzeni liniowej. Dzięki temu możemy wykorzystać do naszego celu narzędzia z algebry liniowej i (w bardzo podstawowym zakresie) topologii.

Definicja 4.10. *Dla pewnego zbioru S składającego się z k struktur (nazwijmy je* strukturami bazowymi*) niech $\mathcal{S}$ będzie zbiorem wszystkich struktur reprezentowanych jako sumę pewnych elementów S, tzn. $\mathcal{S} = span^{\mathbb{N}}(S)$.*

[3] Przypomnijmy, że k jest równe liczbie bazowych zapytań, $k = |W|$.



Rysunek 4.1: (Example 4.14) Przykład, dla którego M_W jest osobliwa (w_1 i w_2 są strukturami nad sygnaturą o dwóch relacjach binarnych - zaznaczonych na obrazku jako zielona i czerwona).

Definicja 4.11. *Dla zbioru struktur $S = \{s_1, ..., s_{|S|}\}$ definiujemy jego* macierz ewaluacji $M_S \in \mathbb{R}^{k \times |S|}$ *wzorem $M_S(i,j) = w_i(s_j)$.*

Innymi słowy, wartość w polu (i,j) macierzy M_S jest równa liczbie homomorfizmów z w_i do s_j.

Definicja 4.12. *Zbiór S jest* dobry *gdy $|S| = k$ oraz macierz M_S jest nieosobliwa.*

Naszym celem

Uwaga 4.13. *Zbiór W, składający się z k zapytań, także jest zbiorem k struktur. Czy możemy przyjąć W jako S zbiór struktur bazowych? Poniższy przykład pokazuje, że M_W nie zawsze jest* dobry*:*

Przykład 4.14. *Niech W składa się z w_1 oraz w_2 takich jak na rysunku 4.1, wtedy:*

$$M_W = \begin{array}{cc} & \begin{array}{cc} w_1 & w_2 \end{array} \\ \begin{array}{c} w_1 \\ w_2 \end{array} & \left[\begin{array}{cc} 2 & 4 \\ 1 & 2 \end{array}\right] \end{array}$$

Zatem M_W jest osobliwa.

Resztę dowodu lematu 4.8 przedstawimy dowodząc dwa następujące lematy pomocnicze:

Lemat 4.15. *Istnieje* dobry *zbiór struktur bazowych S.*

Lemat 4.16. *Jeśli $\vec{q} \notin span\{\vec{v} \mid v \in V\}$ i S jest* dobrym *zbiorem struktur bazowych, to istnieją $D, D' \in \mathcal{S}$ spełniające warunki (A) i (B) określone powyżej.*

Uwaga 4.17. *Założenie, że S jest* dobrym *zbiorem struktur bazowych (tzn. że M_S jest nieosobliwą macierzą) jest istotne. Przyjrzyjmy się dalej przykładowi wprowadzonemu na rysunku 4.1.*

Przykład 4.18. *Niech $q = w_1, V_0 = \{w_2\}$, gdzie w_1, w_2 są takie, jak na rysunku 4.1. Wtedy, ponieważ $w_1 \subseteq_{set} w_2$, $V = V_0$, zatem $W = \{w_1, w_2\}$. Załóżmy, że przyjmiemy $S = \{w_1, w_2\}$. Ponieważ $\vec{q} = (1,0)$ oraz $\vec{v} = (0,1)$, $q \not in\{v \mid v \in V\}$, zatem z naszego głównego lematu wynika $V \not\xrightarrow{bag} q$. Powinniśmy być zatem w stanie znaleźć odpowiedni kontrprzykład D, D'. Przeanalizujmy, sytuację, gdybyśmy przyjęli*

$S = W$. Wtedy dowolna bazy $D \in \mathcal{S}$ jest postaci $D = kw_1 + lw_2$, gdzie $k, l \in \mathbb{N}$, więc $q(D) = 2k+l$ (patrz macierz w przykładzie 4.14) oraz $v(D) = 4k+2l$. Zatem $q(D) = v(D)/2$, więc jeśli dla pewnych $D, D' \in \mathcal{S}$ $v(D) = v(D')$, to również $q(D) = q(D')$. Zwróćmy uwagę, że przyczyną tego, że nie potrafimy znaleźć kontrprzykładu jest fakt, że w macierzy ewaluacji M_S jedna z kolumn jest krotnością drugiej. Na szczęście, jak pokazuje lemat 4.15, zawsze możliwy jest również dobór S tak, by macierz M_S była nieosobliwa, czyli by kolumny M_S były liniowo niezależne.

4.3.2 Poszukiwanie dobrego S - dowód lematu 4.15

Nasz dowód opiera się na lemacie zdefiniowanym w pracy [6] i udowodninym, jak wskazuje [3], w pracy [8], która niestety nie jest łatwo dostępna.

Lemat 4.19. *Dwie struktury G, G' są izomorficzne wtedy i tylko wtedy, gdy dla każdej struktury H, $|hom(H, G)| = |hom(H, G')|$.*

Dowód tego lematu jest analogiczny, do dowodu następującego lematu, który można przeczytać w [12].

Lemat 4.20. *Dwie struktury G, G' są izomorficzne wtedy i tylko wtedy, gdy dla każdej struktury H, $|hom(G, H)| = |hom(G', H)|$.*

Skonstruujemy dobry S w następujących 3 krokach:

Krok 1. $S^{(1)} = \{s_1^{(1)}, ..., s_m^{(1)}\}$ jest dowolnym skończonym zbiorem takim, że:

$$\forall w \neq w' \in W \; \exists i \leq m \quad |hom(w, s_i^{(1)})| \neq |hom(w', s_i^{(1)})|$$

Wiemy, że taki zbiór istnieje korzystając z lematu 4.19. Rzeczywiście, ponieważ elementy W są parami nieizomorficzne, wystarczy wziąć po jednej strukturze dla każdej pary $w \neq w' \in W$, która będzie je różnicować.

W dwóch kolejnych krokach będziemy konstruować dobre S używając mnożenia i dodawania struktur. Z lematu 1.4 takie dodawanie/mnożenie struktur odpowiada dodawaniu/mnożeniu odpowiadających im kolumn w macierzy ewaluacji. Zatem ta sekcja będzie mówić bardziej o prostej algebrze liniowej, niż o zliczaniu homomorfizmów.

Krok 2. Niech $T \in \mathbb{N}$ będzie liczbą większą niż wszystkie wartości $M_{S^{(1)}}$. Wtedy zbiór $S^{(2)}$ będzie się składał z jednej struktury określonej jako:

$$s^{(2)} = \sum_{i=1}^{m} T^i s_i^{(1)}$$

Obserwacja 4.21. *Załóżmy, że $w, w' \in W$ i $w \neq w'$. Wtedy $|hom(w, s^{(2)})| \neq |hom(w', s^{(2)})|$.*

Dowód. Z naszego ulubionego lematu 1.4 otrzymujemy:

$$|hom(w, s^{(2)})| = |hom(w, \sum_{i=1}^{m} T^i s_i^{(1)})| = \sum_{i=1}^{m} T^i |hom(w, s_i^{(1)})|$$

Podobnie $|hom(w', s^{(2)})| = \sum_{i=1}^{m} T^i |hom(w', s_i^{(1)})|$. Zwróćmy uwagę, że sekwencja

$$|hom(w, s_m^{(1)})|; |hom(w, s_{m-1}^{(1)})|; \ldots |hom(w, s_1^{(1)})|; 0$$

jest reprezentacją liczby $|hom(w, s^{(2)})|$ w systemie pozycyjnym o podstawie T. Te dwie reprezentacje są różne ponieważ dla pewnego $i \leq m$, $|hom(w, s_i^{(1)})| \neq |hom(w', s_i^{(1)})|$ (patrz krok 1). Zatem liczby odpowiadające tym reprezentacją też są różne. □

Krok 3. Niech $S^{(3)} = \{s_1^{(3)}, ..., s_k^{(3)}\}$, gdzie $s_i^{(3)} = \left(s^{(2)}\right)^{i-1}$. Pokażemy, że macierz $M_{S^{(3)}}$ jest nieosobliwa.

Przypomnijmy, że $M_{S^{(3)}}(i, j) = |hom(w_i, s_j^{(3)})|$ oraz zauważmy:

$$|hom(w_i, s_j^{(3)})| = |hom(w_i, \left(s^{(2)}\right)^{j-1})| = |hom(w_i, s^{(2)})|^{j-1}$$

Pozostaje nam udowodnić następujący prosty lemat o macierzach:

Lemat 4.22. *Niech $a_1, ..., a_k$ będą parami różnymi liczbami rzeczywistymi. Wtedy macierz $A \in \mathbb{R}^{k \times k}$ określona jako $A(i, j) = a_i^{j-1}$ jest nieosobliwa.*

Dowód. Pokażemy, że kolumny macierzy A są liniowo niezależne. Weźmy $\alpha_1, ..., \alpha_k \in \mathbb{R}$ takie, że dla każdego $i \in \{1, \ldots, k\}$, $\sum_{j=1}^{k} \alpha_j A(i, j) = 0$. Wykażemy, że wtedy $\alpha_1 = \cdots = \alpha_k = 0$. Zdefiniujmy wielomian $P(X) = \alpha_1 + \alpha_2 X + ... + \alpha_k X^{k-1}$. Wtedy $\sum_{j=1}^{k} \alpha_j A(i, j) = \sum_{j=1}^{k} \alpha_j a_i^{j-1} = P(a_i)$. Ponieważ $a_1, ..., a_k$ są parami różne, wiemy, że P ma przynajmniej k miejsc zerowych. Jednak stopień P wynosi co najwyżej $k - 1$, więc $P = 0$, czyli wszystkie jego współczynniki są równe 0. □

Pokazaliśmy zatem, że istnieje zbiór struktur bazowych $S = S^{(3)}$, który jest *dobry.*

4.3.3 Wskazanie kontrprzykładu - dowód lematu 4.16

Od teraz zakładamy, że S jest *dobrym* zbiorem struktur bazowych. Ponadto przyjmujemy $M = M_S$.

Jak już można zauważyć, przestrzenie liniowe są kluczowe w zrozumieniu głównych mechanizmów dowodu twierdzenia 3. Wprowadźmy zatem wektorową reprezentację struktur i zobaczmy, jakie własności są z nią związane:

Definicja 4.23. *Dla $s = \sum_{i=1}^{k} a_i s_i$ definiujemy* $\vec{s} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$

Definicja 4.24. *1. Niech* $\vec{u}, \vec{v} \in \mathbb{R}^k$. *Wtedy* $\vec{u} \circ \vec{v} = \begin{bmatrix} \vec{u}(1)\vec{v}(1) \\ \vdots \\ \vec{u}(k)\vec{v}(k) \end{bmatrix}$

2. Niech $\vec{u}, \vec{v} \in \mathbb{R}^k_{\geq 0}$. *Wtedy* $\vec{u} \,\male\, \vec{v} = \prod_{i=1}^k \vec{u}(i)^{\vec{v}(i)}$.

3. Niech $t \in \mathbb{R}_+, \vec{u} \in \mathbb{R}^k$. *Wtedy* $t^{\vec{u}} = \begin{bmatrix} t^{\vec{u}(1)} \\ \vdots \\ t^{\vec{u}(k)} \end{bmatrix}$

Obserwacja 4.25. *1.* $(\vec{u} \circ \vec{v}) \,\male\, \vec{w} = (\vec{u} \,\male\, \vec{w})(\vec{v} \,\male\, \vec{w})$

2. $t^{\vec{u}} \,\male\, \vec{v} = t^{\langle \vec{u}, \vec{v} \rangle}$

Lemat 4.26. *Niech* $v \in V_q, s \in \mathcal{S}$. *Wtedy* $v(s) = (M\vec{s}) \,\male\, \vec{v}$.

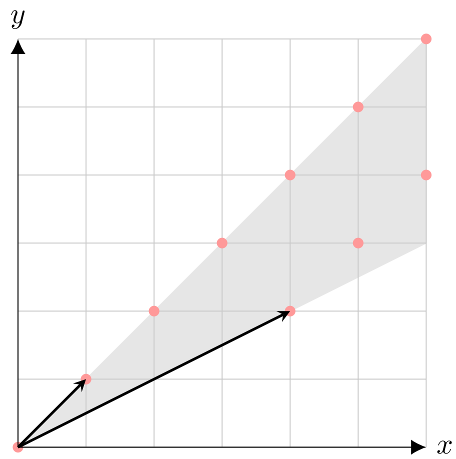
Dowód.

$$\begin{aligned}
(M\vec{s})(i) &= \sum_{j=1}^{k} M(i,j)\vec{s}(j) \\
&= \sum_{j=1}^{k} w_i(s_j)\vec{s}(j) && \text{(z definicji } M\text{)} \\
&= w_i\left(\sum_{j=1}^{k} \vec{s}(j)s_j\right) && \text{(z lematu 1.4)} \\
&= w_i(s) && \text{(z definicji } \vec{s}\text{)}
\end{aligned}$$

$$\begin{aligned}
(M\vec{s}) \,\male\, \vec{v} &= \prod_{i=1}^{k} (M\vec{s})(i)^{\vec{v}(i)} \\
&= \prod_{i=1}^{k} w_i(s)^{\vec{v}(i)} \\
&= \left(\sum_{i=1}^{k} \vec{v}(i)w_i\right)(s) && \text{(z lematu 1.4)}
\end{aligned}$$

□

Jak dotąd określiliśmy dwie k-wymiarowe przestrzenie liniowe, które pomagają nam zrozumieć zachowanie rozważanych obiektów: przestrzeń zapytań koniunkcyjnych z bazą W oraz przestrzeń struktur z bazą S. Nadszedł czas na wprowadzenie trzeciej przestrzeni. Weźmy dowolną strukturę D. Z punktu widzenia danej instancji problemu determinacji, nie interesuje nas jej dokładny kształt, a jedynie odpowiedzi



Rysunek 4.2: (Przykład 4.28) Współrzędna x reprezentuje odpowiedź na w_1, a współrzędna y - na w_2. Czerwone kropki oznaczają punkty z $\mathcal{P}$, natomiast szary obszar to $\mathcal{C}$. Strzałki pokazują wektory kolumnowe M_S.

na bazowe zapytania $w_1, w_2, ..., w_k$ (ponieważ one jednoznacznie wyznaczają odpowiedzi na wszystkie zapytania z V_q). Możemy je reprezentować jako wektor w przestrzeni k-wymiarowej. Będzie to przestrzeń potencjalnych odpowiedzi na zapytania bazowe.

Definicja 4.27. *1.* $\mathcal{P} = \{M\vec{s} \mid s \in \mathcal{S}\} = \{M\vec{u} \mid \vec{u} \in \mathbb{N}^k\}$. *Innymi słowy,* $\mathcal{P}$ *jest zbiorem wszystkich wektorów odpowiedzi na strukturach ze zbioru* $\mathcal{S}$.

2. $\mathcal{C} = span^{\mathbb{R}_{\geq 0}}\{M\vec{s} \mid s \in S\} = span^{\mathbb{R}_{\geq 0}}\{Me_i \mid i \in \{1, \dots, k\}\}$. *Czyli* $\mathcal{C}$ *jest stożkiem wypukłym (ang.* convex cone*) generowanym przez przez wektory bazowe przemnożone przez* M *(równoważnie: generowanym przez* $\mathcal{P}$*),*

Widzimy zatem, że $\mathcal{P}$ jest podzbiorem $\mathcal{C}$, co więcej $\mathcal{P} \subseteq \mathcal{C} \cap \mathbb{N}^k$. Warto zwrócić uwagę, że niekoniecznie $\mathcal{P} = \mathcal{C} \cap \mathbb{N}^k$.

Przykład 4.28. *Niech* w_1, w_2 *będą takie, jak na rysunku 4.1 (oraz w przykładzie 4.14). Niech* s_1 *będzie, grafem składającym się z pojedynczego wierzchołka z czerwoną i zieloną pętelką, natomiast* $s_2 = w_2$*. Wtedy:*

$$M_S = \begin{array}{c c} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} w_1 \\ w_2 \end{array} & \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix} \end{array}$$

Wtedy $\mathcal{C}$ *i* $\mathcal{P}$ *są takie, jak na rysunku* 4.2*. Zwróć uwagę, że teraz* M_S *jest nieosobliwa. To może mieć związek z tym, że szary obszar na rysunku* 4.2 *ma niepuste wnętrze.*

Zbliża się rozwiązanie akcji - miejsce, gdzie seria zagadkowych lematów i warunków zacznie się ze sobą łączyć prowadząc nas do ukończenia dowodu.

Na początek, zauważmy, że rozważanie stożka $\mathcal{C}$, a dokładniej jego podzbioru składającego się z wektorów o współrzędnych wymiernych, ma sens:

Lemat 4.29. *Niech* $\vec{u} \in \mathcal{C} \cap \mathbb{Q}^k$*. Wtedy istnieje* $c \in \mathbb{N}_+$ *takie, że* $c\vec{u} \in \mathcal{P}$*.*

Dowód. M jest nieosobliwa[4] i $M \in \mathbb{Q}^{k\times k}$. Zatem istnieje $M^{-1} \in \mathbb{Q}^{k\times k}$. Niech $\vec{\alpha} = M^{-1}\vec{u}$. Ponieważ $\vec{u} \in \mathcal{C} \cap \mathbb{Q}^k$ otrzymujemy $\vec{\alpha} \in \mathcal{C} \cap \mathbb{Q}^k$. Jako, że $\vec{u} \in \{M\vec{v} \mid \vec{v} \in \mathbb{R}^k_{\geq 0}\}$ wiemy, że $\vec{\alpha} \in \mathbb{R}^k_{\geq 0}$. Zatem istnieje[5] $c \in \mathbb{N}_+$ takie, że $c\vec{\alpha} \in \mathbb{N}_+$. Ostatecznie, $c\vec{u} = c(M\vec{\alpha}) = M(c\vec{\alpha}) \in \mathcal{P}$. □

Lemat 4.30. *Istnieją $\vec{p}, \vec{p'} \in \mathcal{C} \cap \mathbb{Q}^k$ takie, że:*

1. $\forall v \in V\ \vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v} = \vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v}$

2. $\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{q} \neq \vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{q}$

Zanim przystąpimy do dowodu powyższego lematu, zauważmy, że **kończy on dowód lematu 4.16**. Rzeczywiście, jeśli znajdziemy $\vec{p}, \vec{p'}$ takie, jak w lemacie 4.30, to z lematu 4.29 znajdziemy $c, c' \in \mathbb{N}_+$ takie, że $c\vec{p}, c'\vec{p'} \in \mathcal{P}$. Wtedy także $cc'\vec{p}, cc'\vec{p'} \in \mathcal{P}$. Niech $\vec{cc'} = \begin{bmatrix} cc' \\ \vdots \\ cc' \end{bmatrix}$. Wtedy dla $v \in V_q$ mamy:

$$
\begin{aligned}
&(cc'\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v}) - (cc'\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v}) = \\
&= (\vec{cc'} \circlearrowleft\!\!\!\!\nearrow \vec{v})(\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v}) - (\vec{cc'} \circlearrowleft\!\!\!\!\nearrow \vec{v})(\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v}) \qquad \text{(z obserwacji 4.25)} \\
&= (\vec{cc'} \circlearrowleft\!\!\!\!\nearrow \vec{v})((\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v}) - (\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v}))
\end{aligned}
$$

Skoro $cc' > 0$, także $(\vec{cc'} \circlearrowleft\!\!\!\!\nearrow \vec{v}) > 0$ i tak otrzymujemy:

$$(cc'\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v}) - (cc'\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v}) = 0 \iff (\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v}) - (\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v}) = 0$$

Teraz bierzemy $s, s' \in \mathcal{S}$ takie, że $M\vec{s} = cc'\vec{p}, M\vec{s'} = cc'\vec{p'}$. Z lematu 4.26 mamy:

1. dla $v \in V$, $v(s) = cc'\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{v} = cc'\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{v} = v(s')$

2. $q(s) = cc'\vec{p} \circlearrowleft\!\!\!\!\nearrow \vec{q} \neq cc'\vec{p'} \circlearrowleft\!\!\!\!\nearrow \vec{q} = q(s')$

Zatem $D = s, D' = s'$ są strukturami obiecanymi w lemacie 4.16. Do kompletności dowodu pozostaje nam tylko:

Dowód lematu 4.30. Weźmy $\vec{z_0} \in \mathbb{Q}^k$ spełniające: (1) $\forall v \in V\ \langle \vec{z_0}, \vec{v} \rangle = 0$ i (2) $\langle \vec{z_0}, \vec{q} \rangle \neq 0$. Takie $\vec{z_0}$ istnieje z faktu 1.5 oraz założenia, że $\vec{q} \notin \mathrm{span}\{\vec{v} \mid v \in V\}$. Weźmy $d \in \mathbb{N}_+$ takie, że $d\vec{z_0} \in$ i niech $\vec{z} = d\vec{z_0}$. Oczywiście $\vec{z}$ spełnia też warunki (1) i (2).

[4] Wreszcie udało nam się skorzystać z nieosobliwości M. Ale to lemat 4.30 będzie głównym miejscem, gdzie odgrywa ona kluczową rolę.

[5] Wystarczy wziąć dowolną dodatnią wspólną wielokrotność mianowników na wszystkich współrzędnych $\vec{\alpha}$.

Niech $\vec{p} \in \mathcal{C} \cap \mathbb{Q}^k, r > 0$ będą takie, jak we wniosku 1.8, tzn. takie, że kula o promieniu r i środku w $\vec{p}$ zawiera się w $\mathcal{C}$. Mamy już $\vec{p}$ - znajdziemy jeszcze $\vec{p'}$ właśnie w tej kuli.

Lemat 4.31. *Istnieje $t \in \mathbb{R}_+ \setminus \{1\}$ takie, że $t^{\vec{z}} \circ \vec{p} \in \mathcal{C} \cap \mathbb{Q}^k$.*

Dowód. Zauważ, że funkcja $t \mapsto t^{\vec{z}} \circ \vec{p}$ jest ciągła i przenosi 1 na $\vec{p}$. Zatem istnieje $\delta > 0$ taka, że:

$$\forall t \in (1-\delta, 1+\delta) \ \ \|t^{\vec{z}} \circ \vec{p} - \vec{p}\| < r$$

Wystarczy teraz[6] wziąć dowolne $t \neq 1$ ze zbioru $(1-\delta, 1+\delta) \cap \mathbb{Q}$. □

Niech $\vec{p'} = t^{\vec{z}} \circ p$ gdzie t jest jak w lemacie 4.31. Z obserwacji 4.25:

1. Dla $v \in V$, $\vec{p'} \male \vec{v} = (t^{\vec{z}} \circ p) \male \vec{v} = (t^{\vec{z}} \male \vec{v})(\vec{p} \male \vec{v}) = t^{\langle \vec{z}, \vec{v} \rangle}(\vec{p} \male \vec{v}) = \vec{p} \male \vec{v}$
2. $\vec{p'} \male \vec{q} = (t^{\vec{z}} \circ p) \male \vec{q} = (t^{\vec{z}} \male \vec{q})(\vec{p} \male \vec{q}) = t^{\langle \vec{z}, \vec{q} \rangle}(\vec{p} \male \vec{q}) \neq \vec{p} \male \vec{q}$

To kończy dowód lematu 4.30, lematu 4.16 oraz twierdzenia 3. □

4.4 Uwaga o homomorfizmach struktur

Zauważmy, że powyższe rozważania można opowiedzieć zupełnie nie mówiąc o zapytaniach koniunkcyjnych, a jedynie o liczbie homomorfizmów struktur. Istotnie, dla boole'owskiego CQ q i struktury D, mamy $q(D) = |hom(F_q, D)|$. Oczywiście też każdą strukturę możemy przedstawić jako *frozen body* pewnego zapytania. W związku z tym, z lematu 4.8 wynika bezpośrednio następujący wniosek

Wniosek 4.32. *Niech $H_1, ..., H_n, H$ będą pewnymi parami nieizomorficznymi strukturami* ***spójnymi****. Wtedy nie istnieje funkcja $f : \mathbb{N}^n \to \mathbb{N}$ taka, że dla każdej struktury G,*

$$|hom(H, G)| = f(|hom(H_1, G)|, ..., |hom(H_n, G)|)$$

Dowód. Niech $v_1, ..., v_n, q$ będą boole'owskimi CQ takimi, że ich *frozen body* to, odpowiednio, $H_1, ..., H_n, H$. Niech $V = \{v_1, ..., v_n\}$. Definiujemy wektory $\vec{v}_1, ..., \vec{v}_n, \vec{q}$ tak jak w definicji 4.6. Wtedy $\vec{v}_1, ..., \vec{v}_n, \vec{q}$ są wektorami z bazy standardowej. Co więcej, ponieważ odpowiadające im zapytania $v_1, ..., v_n, q$ są parami nieizomorficzne, wektory te są parami różne. Zatem $\vec{q} \notin \mathrm{span}\{\vec{v}_1, ..., \vec{v}_n\}$, więc z lematu 4.8, $V \not\xrightarrow{\text{bag}} q$. Wynika z tego, że nie istnieje funkcja f opisana we wniosku 4.32 □

Stwierdzenie tożsame z powyższym wnioskiem pojawia się w pracy [7] (choć jako luźne sformułowanie we wstępie, nie jako precyzyjnie określone twierdzenie).

[6] Tutaj korzystamy z faktu, że $\vec{z} \in^k$, w przeciwnym wypadku $t^{\vec{z}}$ nie miałoby współrzędnych wymiernych.

Jest ono później przywoływane między innymi w [13]. Co ciekawe jednak, nie zostało ono w tej pracy (ani w żadnej innej, do których udało nam się dotrzeć) udowodnione! Pojawiło się co prawda, z dowodem, pewne twierdzenie, które przez nieuwagę można utożsamić z wnioskiem 4.32. Przytoczymy je i wyjaśnimy, dlaczego wcale nie implikuje ono tego wniosku.

4.4.1 Gęstości homomorfizmów

Ustalmy pewne parami nieizomorficzne grafy[7] spójne $H_1, ..., H_n$.

Definicja 4.33.

1. *Niech G, H będą dowolnymi grafami. Wtedy $t_H(G) = |hom(H, G)|/|G|^{|H|}$, tzn. $t_H(G)$ jest prawdopodobieństwem, że losowo wybrane odwzorowanie z H w G będzie homomorfizmem.*

2. *$S = \{\langle t_{H_1}(G), ..., t_{H_n}(G)\rangle \mid G$ - graf $\}$. Ponadto $\overline{S}$ jest domknięciem topologicznym S, tzn.*

$$\overline{S} = \left\{ v \in \mathbb{R}^n \mid \exists (v_k)_{k \in \mathbb{N}} \in S \quad v_k \xrightarrow[k \to \infty]{} v \right\}$$

Oczywiście $S \subseteq \overline{S}$, ale $S \neq \overline{S}$ - wystarczy zauważyć, że $S \subseteq \mathbb{Q}^n$, natomiast $\overline{S} \nsubseteq \mathbb{Q}^n$, co wynika z faktu, który zaraz przytoczymy.

Głównym twierdzeniem udowodnionym w [7] jest

Fakt 4.34. *Istnieje otwarta n-wymiarowa kula B taka, że $B \subseteq \overline{S}$.*

Można ten fakt wyrazić również w następujący sposób: *Istnieje otwarta n-wymiarowa kula B taka, że zbiór $S \cap B$ jest gęsty w B.*

Następnie w książce [13] stwierdzono, że powyższy fakt implikuje, że nie ma żadnych funkcyjnych zależności pomiędzy $t_{H_1}, ..., t_{H_n}$, co można rozumieć tak, że

(*) nie istnieje funkcja $f : \mathbb{Q}^{n-1} \to \mathbb{Q}$ taka, że dla każdego grafu G $t_{H_n}(G) = f(t_{H_1}(G), ..., t_{H_{n-1}}(G))$

Gdyby tak było, można by zastosować technikę opisaną w [13] (twierdzenie 5.32) tak, by udowodnić wniosek 4.32, przynajmniej w wersji grafowej. Zwróćmy jednak uwagę, że **fakt 4.34 nie implikuje (*)**.

Stwierdzenie **(*)** mówi o tym, że S nie można rozszerzyć do wykresu pewnej funkcji $f : \mathbb{Q}^{n-1} \to \mathbb{Q}$. Jednak wszystko, co wiemy z faktu 4.34, to to, że S jest gęsty w pewnej kuli B. Oznacza to m. in., że nie można go rozszerzyć do funkcji

[7] Zachowując konwencję z pracy [7], będziemy w tej sekcji mówić o grafach, tzn. strukturach spójnych nad pojedynczą relacją binarną. Jednak równie dobrze moglibyśmy to samo napisać również o strukturach.

f, która byłaby ciągła, nie wyklucza jednak takiego rozszerzenia do funkcji nieciągłej. Istotnie, skonstruowanie funkcji $f : \mathbb{Q}^{n-1} \to \mathbb{Q}$, której wykres będzie gęsty w $\mathbb{R}^n$ jest prostym ćwiczeniem. I choć wydaje się niewiarygodne, by tego typu funkcja mogła naturalnie występować przy gęstościach homomorfizmów, nie zostało to udowodnione, w związku z czym, rozumowanie przedstawione w naszej pracy jest pierwszym znanym nam pełnym dowodem wniosku 4.32.

Rozdział 5

Pozostałe przypadki CQ - otwarte problemy

Jak dotąd udało nam się pokazać nierozstrzygalność determinacji dla unii zapytań koniunkcyjnych (co nie było szczególnie trudne), a także rozstrzygalność w dwóch szczególnych przypadkach CQ. Co można powiedzieć o zapytaniach koniunkcyjnych w ogólności? Nie udało nam się rozwiązać tego problemu, jednak podejrzewamy, że problem determinacji w semantyce multizbiorowej jest rozstrzygalny w przypadku unarnych zapytań koniunkcyjnych (a przynajmniej przy dodatkowym założeniu, że są one spójne) oraz nierozstrzygalny w przypadku zapytań o wyższej arności.

Skąd bierze się taki podział? Okazuje się, że zapytania unarne przechwytują jedynie *lokalne* właściwości struktur, ponadto ich zachowanie w dużej mierze przypomina zachowanie zapytań boole'owskich (dokładniej wyjaśnimy to w lemacie 5.1). Natomiast już zapytania binarne potrafią przechwytywać *nielokalne* właściwości struktur (to znaczy takie, które obejmują wierzchołki oddalone od siebie o dowolną odległość w grafie Gaifmana). Dla przykładu. załóżmy, że rozważamy sygnaturę składającą się z jednej binarnej relacji E. Wtedy odpowiedź na zapytanie $v(x, y) = E(x, y)$ niesie całą informację o strukturze, czyli determinuje wszystkie zapytania. W szczególności, dla każdego $n \in \mathbb{N}$ determinuje zapytanie „czy wierzchołki x, y są połączone ścieżką długości n?".

5.1 Unarne zapytania koniunkcyjne

Lemat 5.1. *Niech Σ będzie sygnaturą relacyjną. Niech V będzie zbiorem spójnych*[1]*, unarnych zapytań koniunkcyjnych oraz niech q będzie spójnym unarnym CQ (wszystkie zapytania są nad sygnaturą Σ). Niech $\Sigma' = \Sigma \cup \{c\}$, gdzie c jest symbolem stałej oraz niech V', q' stanowią boole'owskie zapytania utworzone z V, q poprzez zamianę*

[1] spójnych, tzn. takich, dla których ich graf Gaifamana jest spójny

wszystkich wystąpień jedynej zmiennej wolnej na c. Wtedy

$$V \xrightarrow{bag} q \iff V' \xrightarrow{bag} q'$$

Dowód.

- $V \xrightarrow{\text{bag}} q \Leftarrow V' \xrightarrow{\text{bag}} q'$
  Załóżmy nie wprost, że $V' \xrightarrow{\text{bag}} q'$ i $V \not\xrightarrow{\text{bag}} q$. Wtedy istnieją bazy danych D_1, D_2 nad sygnaturą Σ takie, że $V(D_1) = V(D_2)$ oraz $q(D_1) \neq q(D_2)$. $q(D_1) \neq q(D_2)$ oznacza, że istnieje $a \in D_1, D_2$ taki, że $q(D_1)[a] \neq q(D_2)[a]$. Niech D_1', D_2' będą bazami danych nad sygnaturą Σ' które powstają z D_1, D_2 poprzez przypisanie stałej c do elementu a. Wtedy $V'(D_1') = V'(D_2')$ (ponieważ dla każdego $v \in V$, $v'(D_1') = v(D_1)[a] = v(D_2)[a] = v'(D_2')$) oraz $q'(D_1') \neq q'(D_2')$ (ponieważ $q'(D_1') = q(D_1)[a] \neq q(D_2)[a] = q'(D_2')$. Sprzeczność.

- $V \xrightarrow{\text{bag}} q \Rightarrow V' \xrightarrow{\text{bag}} q'$
  Załóżmy nie wprost, że $V \xrightarrow{\text{bag}} q$ i $V' \not\xrightarrow{\text{bag}} q'$. Wtedy istnieją bazy danych D_1', D_2' nad sygnaturą Σ' takie, że $V'(D_1') = V'(D_2')$ oraz $q'(D_1') \neq q'(D_2')$. Bez straty ogólności przyjmijmy, że $D_1' \cap D_2' = \emptyset$. Niech c_1, c_2 będą elementami odpowiednio D_1', D_2', przypisanymi do symbolu stałej c. Niech D_1, D_2 będą strukturami nad sygnaturą Σ powstałymi z D_1', D_2' (w których po prostu zapominamy o stałej). Oczywiście $c_1 \in D_1, c_2 \in D_2$, za to przestały one być elementami wyróżnionymi jako stałe.

  Niech $E_1 = D_1 \cup D_2$ oraz niech E_2 będzie strukturą powstałą z E_1 poprzez zamianę c_1, c_2. To znaczy, że następująca funkcja $f : E_1 \to E_2$ jest izomorfizmem:

$$f(x) = \begin{cases} c_2 \text{ gdy } x = c_1 \\ c_1 \text{ gdy } x = c_2 \\ x \text{ w p. p.} \end{cases}$$

  Pokażemy, że $V(E_1) = V(E_2)$ i $q(E_1) \neq q(E_2)$, co przeczy, że $V \xrightarrow{\text{bag}} q$. Zauważmy, że ponieważ f jest izomorfizmem, dla dowolnego zapytania unarnego $\varphi(x)$ oraz dla $a \in E_1$ zachodzi:

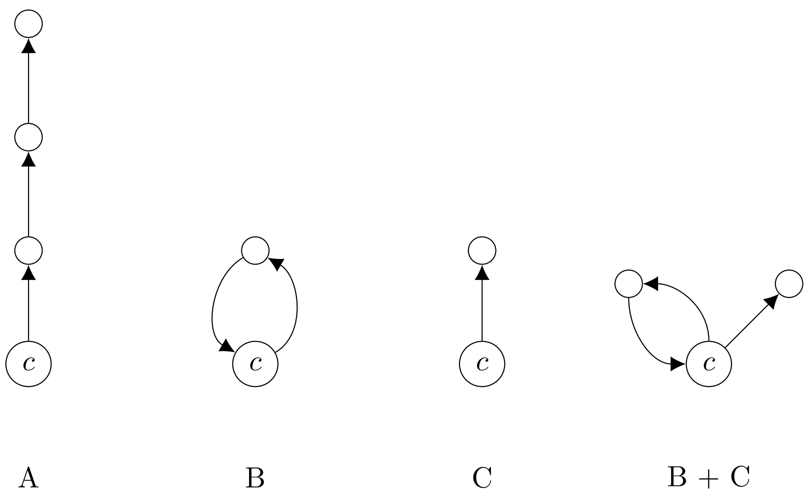
$$\varphi(E_1)[a] = \varphi(E_2)[f(a)]$$

  Wykorzystamy ten fakt, by pokazać $V(E_1) = V(E_2)$. Niech $v \in V$. Oczywiście, ponieważ dla $x \neq c_1, c_2$, $f(x) = x$, wiemy, że $v(E_1)[x] = v(E_2)[x]$. Ponadto, $v(E_1)[c_1] = v'(D_1') = v'(D_2') = v(E_1)[c_2]$. Równość $v(E_1)[c_i] = v'(D_i)$ wynika z założenia, że v jest zapytaniem spójnym, natomiast E_1 jest sumą rozłączną D_1 oraz D_2 (zatem wszystkie zmienne występujące w zapytaniu muszą przejść na elementy tej samej spójnej składowej - albo zawartej w D_1 albo w D_2). Zatem $v(E_1)[c_1] = v(E_1)[c_2] = v(E_2)[f(c_2)] = v(E_2)[c_1]$. Analogicznie, $v(E_1)[c_2] = v(E_2)[c_2]$, co kończy dowód, że $v(E_1) = v(E_2)$. Z kolei $q(E_1)[c_1] = q'(D_1') \neq q'(D_2)' = q(E_1)[c_2] = q(E_2)[c_1]$. Zatem $q(E_1) \neq q(E_2)$ - sprzeczność.

□

Czy powyższy lemat oraz twierdzenie 3 (o rozstrzygalności determinacji dla boole'owskich CQ) dowodzą, że również problem determinacji spójnych unarnych CQ jest rozstrzygalny? Niestety nie, ponieważ zapytania V', q' z lematu 5.1 są nad sygnaturą, która zawiera nie tylko relacje, ale także symbol stałej c. Wydaje się to być niewielką różnicą, ale sprawia, że przestaje mieć sens podstawa, na której oparty jest dowód twierdzenia 3: operacja dodawania struktur wymaga przedefiniowania - nie możemy brać rozłącznej sumy struktur, ponieważ wtedy mielibyśmy dwa elementy odpowiadające stałej c.

Najbardziej rozsądną definicją jest taka, że $A + B = A \cup B'$, gdzie B' jest strukturą izomorficzną z B taką, że $A \cap B' = \{c\}$ (czyli innymi słowy, bierzemy sumę rozłączną A, B, a następnie „sklejamy" ich wystąpienia stałej c). Wtedy jednak nie działa już lemat 1.4, a dokładniej własność, że gdy A jest spójne, to $|hom(A, B + C)| = |hom(A, B)| + |hom(A, C)|$.



Przykład 5.2. *Na rysunku powyżej, jak łatwo policzyć* $|hom(A, B)| = 1$, $|hom(A, C)| = 0$*, natomiast* $|hom(A, B + C)| = 2$*. Pojawił się jeden homomorfizm, który część struktury A przenosi na składową pochodzącą z B, a część na składową pochodzącą z C.*

Podjęliśmy pewne próby naprawienia tej luki - jednym z pomysłów było zastąpienie dodawania struktur przez mnożenie, które jest dobrze określone dla struktur ze stałą i spełnia (wraz z dodawaniem określonym wcześniej) wszystkie pozostałe własności opisane w lemacie 1.4. Ta próba jednak zawiodła, ponieważ nie da się w niej w rozsądny sposób nadać sensu punktom o współrzędnych wymiernych w przestrzeni potencjalnych odpowiedzi (patrz lemat 4.29 oraz komentarz po zdefiniowaniu lematu 4.30). Tam korzystaliśmy z faktu, że mając wektor o współrzędnych wymiernych możemy go przemnożyć przez odpowiednią liczbę tak, by otrzymać wektor ze zbioru $\mathcal{P}$ (w którym wszystkie wektory miały współrzędne naturalne). W tym wypadku, zamiast mnożenia musielibyśmy ten wektor podnieść do pewnej potęgi (tzn. podnieść do tej samej potęgi jego wartości na wszystkich współrzędnych). Czasem jednak nie da się tego zrobić tak, by otrzymać wektor o współrzędnych całkowitych. Na przykład, niech $\vec{v} = \left(\frac{1}{2}, \frac{1}{3}\right)$. Wtedy nie istnieje $\alpha \in \mathbb{R}$ taka, że $\left(\frac{1}{2}\right)^\alpha \in \mathbb{N}$ oraz $\left(\frac{1}{3}\right)^\alpha \in \mathbb{N}$, zatem jesteśmy skazani na niepowodzenie przy zastosowaniu triku z lematu 4.29.

Pomimo tych niepowodzeń, podejrzewamy, że problem determinacji boole'owskich zapytań koniunkcyjnych ze stałą (a zatem i spójnych unarnych zapytań koniunkcyjnych) jest rozstrzygalny. Cała trudność w udowodnieniu tego leży w znalezieniu kontrprzykładu determinacji wtedy, gdy nie wynika ona w oczywisty sposób z konstrukcji zapytań (por. lemat 4.8). Wydaje się jednak, że trudność w znalezieniu tego kontrprzykładu wynika z tego, że świat struktur ze stałą jest bardziej chaotyczny pod względem zliczania homomorfizmów niż świat struktur bez stałej, a nie stąd że jest on bardziej ubogi.

Bibliografia

[1] Mahmoud Abo Khamis, Phokion G. Kolaitis, Hung Q. Ngo, and Dan Suciu. Bag query containment and information theory. In *Proceedings of the 39th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, PODS'20, page 95–112, New York, NY, USA, 2020. Association for Computing Machinery.

[2] Foto Afrati. Determinacy and query rewriting for conjunctive queries and views. *Theor. Comput. Sci.*, 412:1005–1021, 03 2011.

[3] Albert Atserias, Phokion G. Kolaitis, and Wei-Lin Wu. On the expressive power of homomorphism counts. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–13. IEEE, 2021.

[4] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Counting graph homomorphisms. In Martin Klazar, Jan Kratochvíl, Martin Loebl, Jiří Matoušek, Pavel Valtr, and Robin Thomas, editors, *Topics in Discrete Mathematics*, pages 315–371, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[5] Ashok K. Chandra and Philip M. Merlin. Optimal implementation of conjunctive queries in relational data bases. In *Proceedings of the Ninth Annual ACM Symposium on Theory of Computing*, STOC '77, page 77–90, New York, NY, USA, 1977. Association for Computing Machinery.

[6] Surajit Chaudhuri and Moshe Y. Vardi. Optimization of real conjunctive queries. In *Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, PODS '93, page 59–70, New York, NY, USA, 1993. Association for Computing Machinery.

[7] Paul Erdős, László Lovász, and Joel Spencer. Strong independence of graphcopy functions, 1979.

[8] S. Fisk. Distinguishing graphs by the number of homomorphisms. *Discuss. Math. Graph Theory*, 15:73–75, 1995.

[9] Grzegorz Głuch, Jerzy Marcinkowski, and Piotr Ostropolski-Nalewaja. Can one escape red chains? regular path queries determinacy is undecidable. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '18, page 492–501, New York, NY, USA, 2018. Association for Computing Machinery.

[10] Tomasz Gogacz and Jerzy Marcinkowski. The hunt for a red spider: Conjunctive query determinacy is undecidable. In *Proceedings of the 2015 30th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, LICS '15, page 281–292, USA, 2015. IEEE Computer Society.

[11] Jaroslaw Kwiecien, Jerzy Marcinkowski, and Piotr Ostropolski-Nalewaja. Determinacy of real conjunctive queries. the boolean case. *CoRR*, abs/2112.12742, 2021.

[12] László Lovász. Operations with structures. *Acta Mathematica Academiae Scientiarum Hungarica*, 18(3-4):321–328, 1967.

[13] László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.

[14] Jerzy Marcinkowski. What Makes a Variant of Query Determinacy (Un)Decidable? (Invited Talk). In Carsten Lutz and Jean Christoph Jung, editors, *23rd International Conference on Database Theory (ICDT 2020)*, volume 155 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:20, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[15] Alan Nash, Luc Segoufin, and Victor Vianu. Determinacy and rewriting of conjunctive queries using views: A progress report. In Thomas Schwentick and Dan Suciu, editors, *Database Theory – ICDT 2007*, pages 59–73, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[16] Luc Segoufin and Victor Vianu. Views and queries: Determinacy and rewriting. In *Proceedings of the Twenty-Fourth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '05, page 49–60, New York, NY, USA, 2005. Association for Computing Machinery.