

Uniwersytet Jagielloński
Wydział Matematyki i Informatyki
INSTYTUT INFORMATYKI
Studia dzienne

Nr indeksu: WMIF/SMP/06/01/02

Lech Duraj

Złożoność obliczeniowa niektórych problemów kafłowania

Opiekun pracy magisterskiej:
prof. dr hab. Paweł Idziak

Opracowano zgodnie z Ustawą o prawie autorskim i prawach pokrewnych z dnia 4 lutego 1994 r. (Dz.U. 1994 nr 24 poz. 83) wraz z nowelizacją z dnia 25 lipca 2003 r. (Dz.U. 2003 nr 166 poz. 1610) oraz z dnia 1 kwietnia 2004 r. (Dz.U. 2004 nr 91 poz. 869)

Kraków 2006

SPIS TREŚCI

Streszczenie	2
1. Wstęp	3
2. Przypadek dwóch kolorów	5
3. Trzy kolory — przypadki NP-zupełne	8
3.1. Bramki	8
3.2. Redukcja	15
3.3. Wnioski	18
4. Trzy kolory — przypadki wielomianowe	21
5. Podsumowanie	23
Podziękowania	23
Literatura	24

Streszczenie

Rozważamy następujący *problem hydraulika*: na prostokątnej planszy składającej się z kwadratowych pól umieszczamy (również kwadratowe) płytki z pewnego ustalonego zestawu, o bokach pomalowanych na różne kolory. Każdą z płytek wolno obracać, posiada więc cztery możliwe stany, nie wolno jednak zmieniać jej położenia. Szukamy odpowiedzi na pytanie, czy da się tak dobrać ustawienia wszystkich płytek, aby każda krawędź styku miała po obu stronach te same kolory boków.

Nazwa problemu pochodzi od jego dwukolorowej wersji: płytki możemy wyobrazić sobie jako łączenie rur (w istocie, problem ten występuje w grze komputerowej *Plumber*), a celem naszym jest takie dopasowanie owych rur, aby instalacja nie miała przecieku.

Znane jest prawie pełne rozwiązanie problemu w przypadku, kiedy możliwe są tylko dwa kolory płytek (nazwijmy je *czarnym* i *białym*). Jeśli weźmiemy wszystkie dostępne czarno-białe płytki, problem ten jest NP-zupełny, jednak dla niektórych mniejszych zestawów daje się rozwiązać w czasie wielomianowym ([1]).

W tej pracy rozważamy zestawy płytek o trzech kolorach (*czarny*, *biały* i *czerwony*), oczywiście te z nich, które nie są nadzbiorami żadnych z NP-zupełnych zestawów dwukolorowych. Modyfikując odpowiednio dowód z [1], pokazujemy twierdzenie, które pozwala rozwiązać problem w wielu istotnych przypadkach: otóż wystarczy, aby zestaw zawierał najprostsze płytki czarno-białe i dawał możliwość konstrukcji gadżetu zwanego *flip-flopem*, aby odpowiedni problem był NP-zupełny. Dowód przebiega przez wielomianową redukcję problemu (1,3)-SAT do problemu obracania płytek.

Rozważamy też niektóre inne zestawy płytek, w szczególności wersję problemu, w której zabraniamy płytek istotnie trójkolorowych. Pokazujemy, jeśli wszystkie trzy podzestawy (czarno-białe, czerwono-białe i czerwono-czarne) są wielomianowe, ten problem również daje się rozwiązać w czasie wielomianowym.

1. Wstęp

Rozważmy następującą grę jednoosobową: na prostokątnej planszy podzielonej na kwadratowe pola ułożone są kwadratowe płytki, jedna na każdym polu. Krawędzie płytek pomalowane są na różne kolory. Miejsca płytek na szachownicy dane są z góry, graczowi wolno jedynie je obracać o 90, 180, lub 270 stopni. Każda z płytek ma zatem cztery możliwe stany. Dodatkowo wyróżniamy jeden z kolorów, który będziemy nazywać **białym**.

Celem gry jest znalezienie na planszy takiego **ustawienia** płytek (formalnie: przyporządkowania każdej płytce jednego ze stanów), aby

- każde dwie sąsiednie płytki miały na wspólnej krawędzi ten sam kolor
- każda krawędź płytki będąca zewnętrzną krawędzią prostokąta miała kolor biały.

Żądane ustawienie będziemy nazywać **poprawnym**.

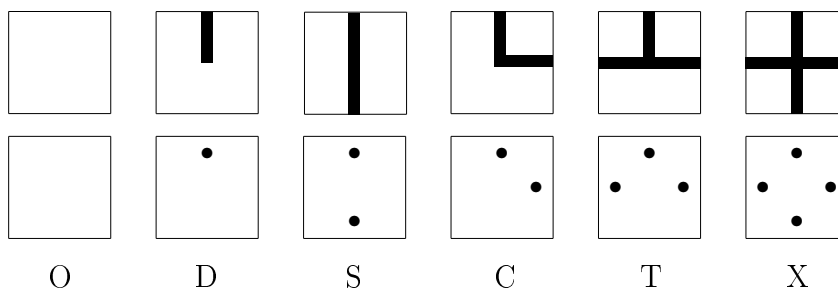
Dla dwukolorowych płytek ta gra jest popularną grą komputerową, znaną jako **Plumber** (hydraulik). Płytki są przedstawiane jako system rur (rysunek 1), dopasowanie kolorów krawędzi odpowiada zaś takiemu połączeniu rur, aby z instalacji nie wyciekała woda.

My będziemy opisywać płytki czwórka liczb określającą kolory krawędzi, kolejno górnej, prawej, dolnej i lewej. Dla płytek o dwóch kolorach użyjemy dodatkowo nazw z [1]:

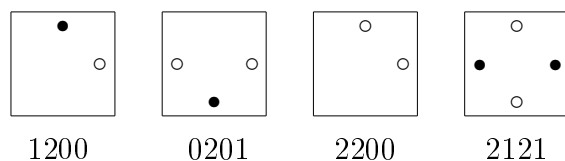
- płytka 0000 to inaczej płytka *O* albo **pusta płytka** (*empty tile*)
- płytka 1000 to płytka *D* (*dead-end tile*)
- płytka 1010 to płytka *S* (*straight tile*)
- płytka 1100 to płytka *C* (*curve tile*)
- płytka 1110 to płytka *T* (*T-junction tile*)
- płytka 1111 to płytka *X* (*crossing tile*)

Oczywiście płytka może mieć więcej niż jedną odpowiadającą jej czwórkę bitów, np. ciągi 0100 i 0010 również opisują płytkę *D*. Konwencja rysunkowa z gry **Plumber** robi się niewygodna dla płytek o trzech lub więcej kolorach, dlatego też w tej pracy używać będziemy innej, pokazanej na rysunkach 1–3.

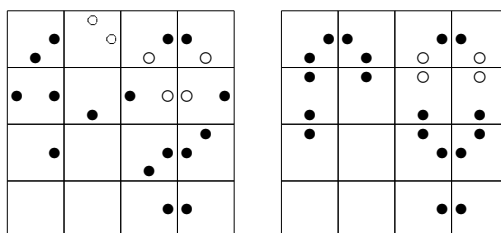
Dla ustalonego zbioru *A* dopuszczalnych rodzajów płytek, rozważmy teraz tak zwany **problem hydraulika**: mając na wejściu planszę *P* płytek z



RYSUNEK 1. Płytki dwukolorowe w obu konwencjach rysunkowych.



RYSUNEK 2. Przykłady płytek trój kolorowych.



RYSUNEK 3. Przykładowa plansza do gry (z lewej) oraz ta sama plansza w poprawnym ustawieniu.

A , chcemy rozstrzygnąć, czy istnieje poprawne ustawienie $\pi(P)$ tej planszy. Jest to problem decyzyjny, w oczywisty sposób rozstrzygalny (wystarczy sprawdzić wszystkie możliwe stany wszystkich płytek — jest ich 4^n , gdzie n jest ilością płytek na wejściu), a nawet w oczywisty sposób w klasie NP — losując stany wszystkich płytek, w wielomianowym czasie weryfikujemy, czy otrzymaliśmy poprawne ustawienie. Będziemy mówili o **dwukolorowym** bądź **trój kolorowym** problemie hydraulika, w zależności od tego, ile różnych kolorów mogą mieć krawędzie płytek z A . Klasa złożoności obliczeniowej problemu hydraulika może oczywiście zależeć od przyjętego zbioru A .

Definicja 1.1. Zbiór (zestaw) rodzajów płytek nazywamy *NP-zupełnym* bądź *wielomianowym*, jeśli związany z nim problem hydraulika należy do klasy złożoności odpowiednio *NPC* i *P*.

2. Przypadek dwóch kolorów

Dwukolorowe zestawy płytek zostały dokładnie rozpatrzone w [1], chociaż jeden z przypadków wciąż pozostaje otwarty.

Twierdzenie 2.1. [1] *Zestaw płytek OCSD jest NP-zupełny.*

Twierdzenie to jest szczególnym przypadkiem twierdzenia 3.14, które udowodnimy w następnym rozdziale.

Twierdzenie 2.2. [1] *Zestawy CSD i CST są NP-zupełne.*

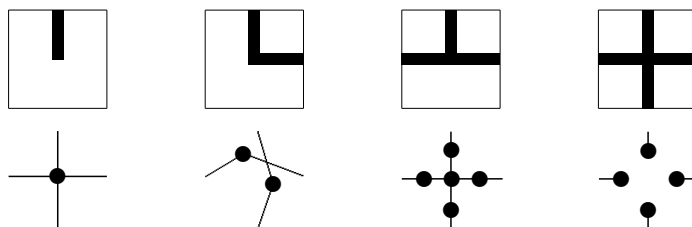
Dowód można znaleźć w [1], jego ideą jest symulacja płytki O za pomocą S i D (przy zestawie CSD), a następnie sprowadzenie problemu CST do CSD przez zamianę koloru czarnego z białym.

Twierdzenie 2.3. [1] *Zestaw płytek OCDTX jest wielomianowy.*

Dowód. Dowód pochodzi z [1], przytoczymy go, aby móc użyć jego modyfikacji w kolejnym twierdzeniu.

Rozważmy planszę P zbudowaną z płytek $OCDTX$. Ideą dowodu jest zbudowanie grafu dwudzielnego G zależnego tylko od planszy P , a następnie pokazanie, że istnieje skojarzenie doskonałe w G wtedy i tylko wtedy, gdy istnieje poprawne ustawienie planszy P .

Dla każdej płytki tworzymy pewien fragment G , jak na rysunku 4.



RYSUNEK 4. Gadzety symulujące płytki D , C , T i X .

Każdemu stykowi dwóch płytek odpowiada pewna krawędź w G , złożona z dwóch części pochodzących z gadżetów płytek. Usuwamy „połówki” krawędzi, czyli krawędzie prowadzące do płytek pustych. Łatwo widać, że powstały graf G jest dwudzielny.

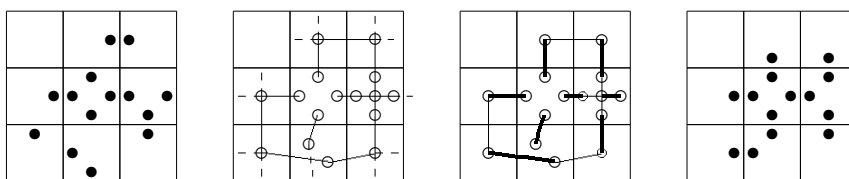
Niech $\pi(P)$ będzie poprawnym ustawieniem płytek. Tworzymy skojarzenie M' w G wybierając do niego te krawędzie, które odpowiadają stykowi czarnych boków płytek. Zauważmy, że:

- płytka D ma dokładnie jeden bok czarny, więc w odpowiadającym jej gadżecie tylko jedna krawędź incydentna z jednym wierzchołkiem gadżetu będzie w skojarzeniu.
- płytka C ma zawsze jeden poziomy i jeden pionowy bok czarny — będą one indukować dwie krawędzie skojarzenia, każda incydentna z jednym z wierzchołków gadżetu.

- płytką X ma wszystkie boki czarne, a zatem wszystkie cztery wierzchołki gadżetu tej płytki będą skojarzone.
- płytką T ma trzy czarne boki, zatem dokładnie trzy spośród skrajnych wierzchołków odpowiedniego gadżetu będą skojarzone.

M' jest zatem poprawnym skojarzeniem w G , co więcej, jedynymi wierzchołkami nieskojarzonymi w M' są po dwa wierzchołki gadżetów płytek T — środkowy i jeden ze skrajnych. Między każdą taką parą istnieje krawędź, jeśli więc wszystkie takie krawędzie dodamy do M' , otrzymamy skojarzenie M w grafie G , które jest doskonałe.

Na odwrót, mając skojarzenie doskonałe M w grafie G potrafimy pomalować boki płytek tak, aby dostać poprawne ustawienie planszy P . Analogicznie jak poprzednio, malujemy na czarno te boki płytek, którym odpowiadają krawędzie z M , zaś na biało pozostałe. W podobny sposób dowodzimy, że po zamianie grafu G na planszę płytki D będą miały po jednym boku czarnym, C - dwa prostopadłe boki, T trzy, zaś X - cztery boki czarne. Co więcej, na każdym styku płytek kolory po obu stronach będą takie same. Powstała plansza jest zatem prawidłowo zorientowaną planszą P .



RYСУNEK 5. Plansza, odpowiadający jej graf, skojarzenie doskonałe oraz poprawne ustawienie płytek

Problem hydraulika dla zestawu płytek $OCDTX$ sprowadza się zatem do pytania, czy w pewnym grafie istnieje skojarzenie doskonałe. Jak wiadomo, algorytm Hopcrofta-Karpa ([2]) rozwiązuje to zagadnienie w czasie $O(\sqrt{|V|}|E|)$, gdzie V jest zbiorem wierzchołków grafu, E zbiorem krawędzi. W tym wypadku, jeśli n jest liczbą płytek na planszy, w grafie znajduje się $O(n)$ wierzchołków i $O(n)$ krawędzi (każda płytkę indukuje w grafie co najwyżej 5 wierzchołków i 8 krawędzi), a zatem czas działania algorytmu wynosił będzie $O(n^{3/2})$. \square

Twierdzenie 2.4. [1] *Zestaw płytek CSOX jest wielomianowy.*

Dowód. Zauważmy, że w przypadku tych czterech płytek kolor górnej i lewej krawędzi wyznacza jedyny możliwy obrót płytki, a zatem i kolor pozostałych krawędzi. Określamy zatem jedyne możliwe ułożenia płytek kolejno rzędami, od lewego górnego rogu, aż dojdziemy do sprzeczności lub prawidłowo zorientujemy wszystkie płytki na planszy. \square

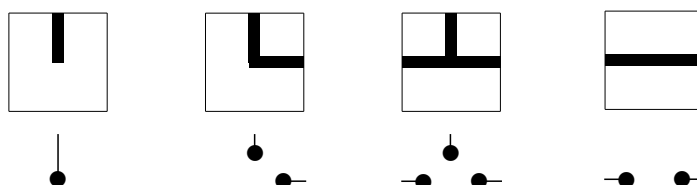
Rozważmy teraz problem dwukolorowego hydraulika z *ustalonym brzegiem*: plansza do gry nie musi być prostokątem, lecz dowolną spójną¹ figurą na

¹spójną — z każdego jej pola na każde inne można dojść przechodząc tylko między sąsiednimi (stykającymi się krawędzią) płytkami.

planszy prostokątnej (dziury są dozwolone), a krawędzie brzegowe nie są zawsze białe, lecz niektóre z nich mogą być czarne. Zauważmy, że równoważnie możemy rozpatrywać standardowy problem hydraulika na prostokątnej planszy, przy czym niektóre płytki zaznaczyć jako nieruchome (nie wolno ich obracać). Płytki nieruchome mogą pochodzić spoza używanego zestawu.

Obserwacja 2.5. *Zestawy płytek z twierdzeń 2.3 i 2.4 pozostają wielomianowe dla problemu hydraulika z ustalonym brzegiem.*

Dowód. Dowód twierdzenia 2.4 oczywiście nie zmienia się, jeśli niektóre krawędzie są już na wejściu ustalone. W dowodzie twierdzenia 2.3 wystarczy wprowadzić nowe gadżety dla ustalonych płytek, analogiczne do gadżetu płytki X (rysunek 6). \square



RYSUNEK 6. Przykłady gadżetów dla nieruchomych płytek, także płytki S .

Obserwacja 2.6. [1] *Dla wszystkich zestawów płytek \mathcal{Z} takich, że $\{D, S, \} \subseteq \mathcal{Z} \subseteq \{O, D, S, T, X\}$ odpowiadające im problemy hydraulika mają tę samą złożoność obliczeniową.*

Nie wiadomo jednak, czy problemy te są w P.

3. Trzy kolory — przypadki NP-zupełne

W tym rozdziale będziemy rozważać trójkolorowe zestawy płytek, które zawierają płytki O , C i D . Innymi słowy, próbujemy odpowiedzieć na pytanie, jakie trójkolorowe płytki wystarczy dołożyć do wielomianowego zestawu płytek czarno-białych, aby problem hydraulika stał się NP-zupełny.

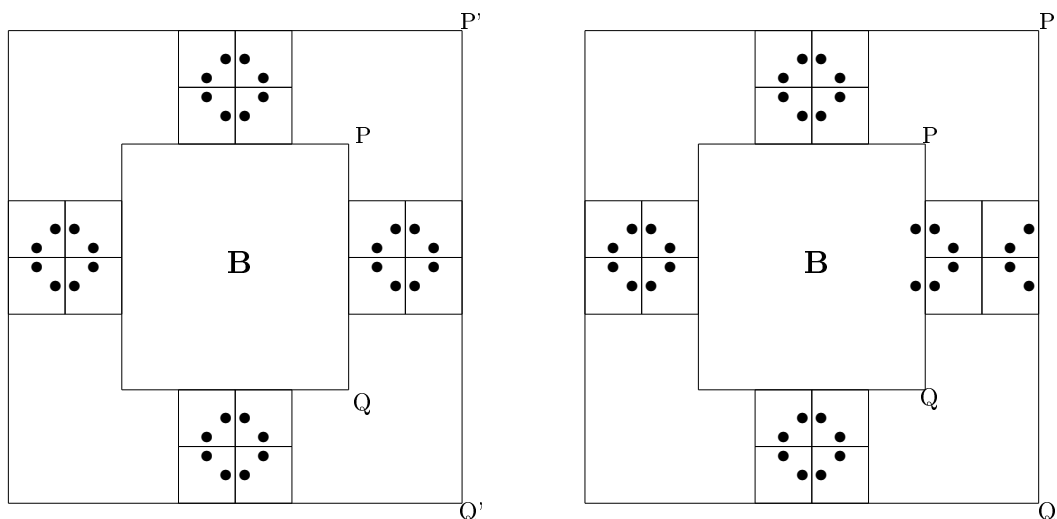
3.1. Bramki.

Definicja 3.1. *Bramką* nazywamy kwadratowy układ płytek, o następujących własnościach:

- (1) bok kwadratu jest liczbą postaci $4k + 2$ dla pewnej liczby naturalnej k (liczbę k nazywamy *rzędem* bramki),
- (2) każdy z zewnętrznych boków kwadratu w dowolnym poprawnym ustawieniu jest albo **nieaktywny** (wszystkie krawędzie płytek na tym boku są białe), albo **aktywny** (wszystkie krawędzie płytek są białe z wyjątkiem dwóch środkowych, czarnych).

Zamiast „bok nieaktywny” czy „aktywny” będziemy mówić też **w stanie 0** i **w stanie 1**. Możemy teraz opisywać możliwe ułożenia bramki czterema cyframi, idąc od górnej krawędzi zgodnie z ruchem wskazówek zegara: na przykład 0010 oznacza bramkę, w której dolna krawędź jest aktywna, zaś pozostałe nieaktywne. Taki czterobitowy ciąg nazwiemy **stanem bramki**. Interesować nas będą bramki, dla których dopuszczalna jest tylko pewna konkretna liczba (zwykle 2 lub 4) ustalonych z góry stanów.

Obserwacja 3.2. *Bramka o stanach s_1, \dots, s_n rzędu $k + 1$ jest zawsze konstruowalna z bramki rzędu k o tych samych stanach.*



RYSUNEK 7. Konstrukcja bramki rzędu $k + 1$ z bramki rzędu k

Dowód. Do środkowych pól każdej krawędzi bramki B dołączamy kwadrat 2×2 z płytek C , resztę przestrzeni wypełniając pustymi płytkami (rysunek

7). Jak widać z rysunku, prawa krawędź $P'Q'$ powstałej bramki ma ten sam stan, co krawędź PQ braki B , podobnie zachowywać się muszą pozostałe krawędzie. Nowa bramka ma zatem te same dopuszczalne stany, co B . \square

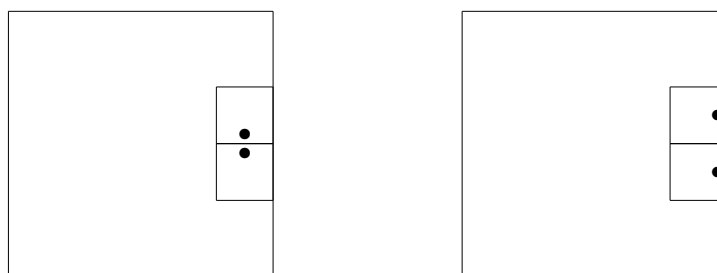
Definicja 3.3. *Flip-flop to bramka, w której dopuszczalne są dokładnie dwa stany: 0101 i 1010.*

Innymi słowy, przeciwległe krawędzie są zawsze w tym samym stanie, a prostopadłe - w różnych.

Definicja 3.4. *Źródło jest bramką, w której prawa krawędź ma dopuszczalne oba stany, a pozostałe trzy są nieaktywne.*

Obserwacja 3.5. *Źródło rzędu k dla każdego $k \geq 0$ jest konstruowalne z czarno-białych płytek O i D .*

Dowód. Umieszczamy dwie płytki D w kwadracie pustych płytek, jak na rysunku 8. \square



RYSUNEK 8. Źródło w obu możliwych stanach.

Możemy wyobrazić sobie źródło jako generator sygnału — w zależności od ustawienia źródła, sygnał pojawia się na jego prawej krawędzi, lub nie.

Definicja 3.6. *Zakręt to bramka, która ma dwa dopuszczalne stany: jednym z nich jest 0000, zaś drugim 1100, 0110, 0011 lub 1001.*

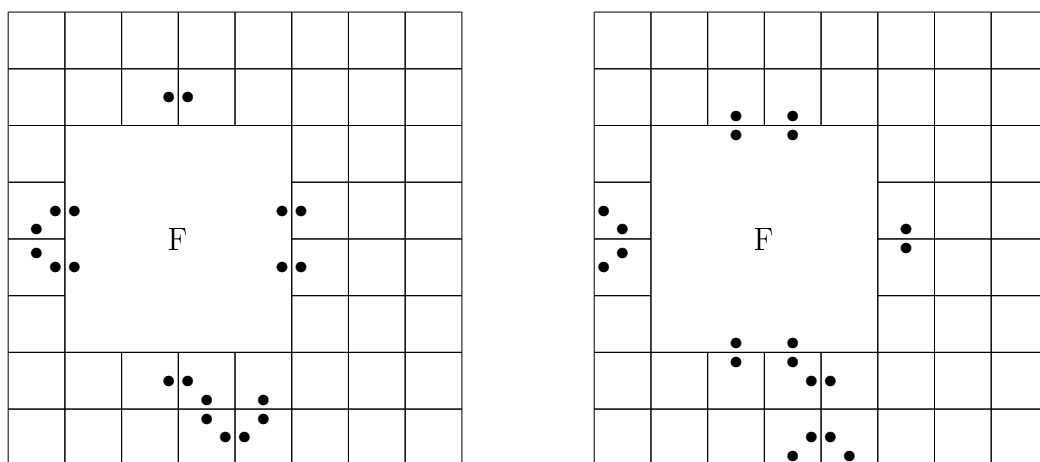
Zakręt wymusza zgodność pewnych dwóch sąsiednich krawędzi — czyli, używając poprzedniej intuicji, przekazuje sygnał z jednej z nich na drugą.

Obserwacja 3.7. *Zakręt rzędu $k + 1$ jest konstruowalny z płytek O , C , D , oraz flip-flopa rzędu k .*

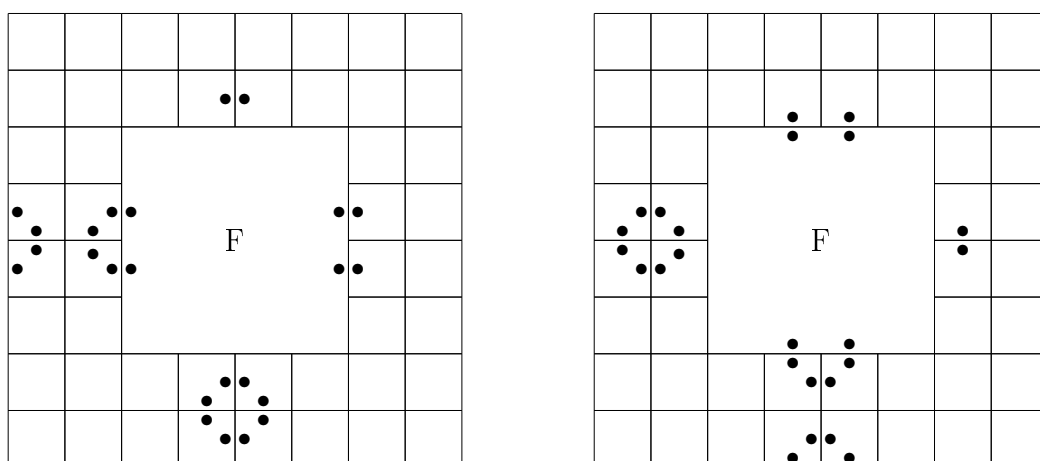
Dowód. Rysunek 9 przedstawia odpowiednią konstrukcję, prostokąty z płytek O, C, D dokładamy na środkach odpowiednich krawędzi flip-flopa F , uzupełniając pustymi płytkami. \square

Definicja 3.8. *Zakręt negowany to bramka, która ma dwa dopuszczalne stany:*

- 0010 i 0001 (zakręt lewo-dół)
- 0001 i 1000 (zakręt lewo-góra)
- 1000 i 0100 (zakręt prawo-góra)



RYSUNEK 9. Konstrukcja zakrętu z bramki flip-flop.



RYSUNEK 10. Konstrukcja zakrętu negowanego z bramki flip-flop.

— 0100 i 0010 (zakręt prawo-dół)

Zakręt negowany sprawia, że pewne dwie sąsiednie krawędzie są w przeciwnych stanach, czyli przekazuje zanegowany sygnał.

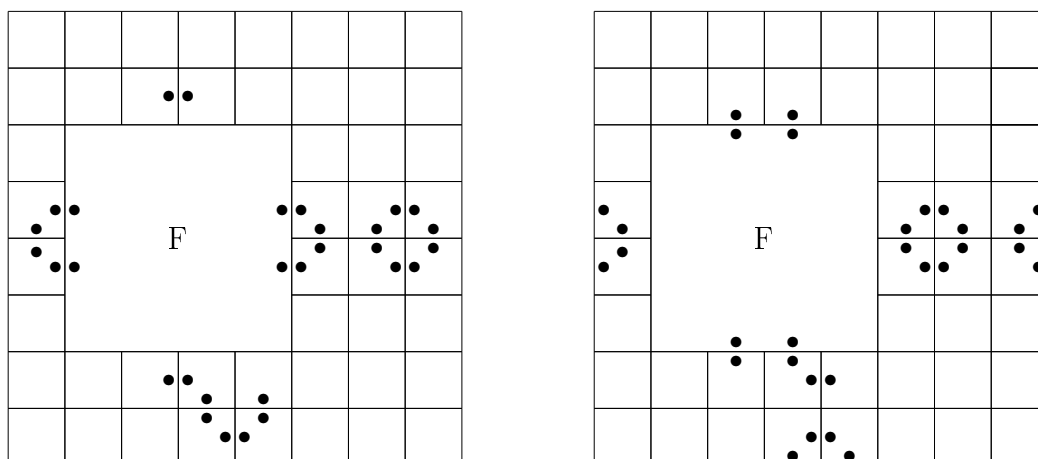
Obserwacja 3.9. Zakręt negowany rzędu $k + 1$ jest konstruowalny z płytek O , C , D i flip-flop'a rzędu k .

Dowód. Analogicznie jak poprzednio, dokładamy układ płytek do każdej krawędzi bramki flip-flop (rysunek 10). \square

Definicja 3.10. *Rozdzielnik* jest branką o następujących dopuszczalnych stanach: 0000 i 0111.

Sygnał z lewej krawędzi przekazywany jest na prawą i dolną.

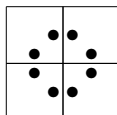
Obserwacja 3.11. Rozdzielnik rzędu $k + 1$ jest konstruowalny z płytek O , C , D , oraz *flip-flopa* rzędu k .



RYSUNEK 11. Konstrukcja rozdzielnika z bramki flip-flop.

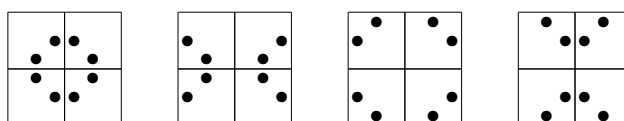
Dowód. Odpowiedni gadżet (rysunek 11) jest bardzo podobny do zakrętu. □

Definicja 3.12. *Przełącznikiem* nazywamy następujący gadżet:

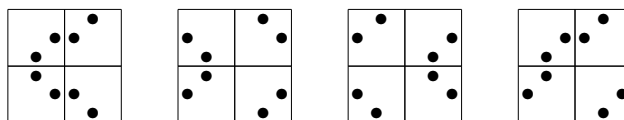


Przełącznik ma cztery dopuszczalne stany: 0000,0101,1010,1111, jednak nie jest bramką, gdyż możliwe są także inne ustawienia płytek (rysunki 12 i 13). Zauważmy jednak, że jeśli wymusimy ten sam kolor na obu bocznych i ten sam na obu górnych krawędziach przełącznika, obie dolne krawędzie, jak również obie boczne krawędzie również będą miały ten sam kolor. Zatem przełącznik sąsiadujący z lewego i górnego boku z prawidłowymi bramkami, sam zachowuje się jak bramka.

Co więcej, możemy do niego podłączać dalsze przełączniki, które w takiej sytuacji można również traktować jak bramki. Gadżet o takiej własności (zachowujący się jak bramka, jeśli prawidłowe kolorowanie na dwóch krawędziach jest wymuszone) nazywać będziemy **pseudobramką**. Z przełącznika



RYSUNEK 12. „Bramkowe” ustawienia przełącznika.

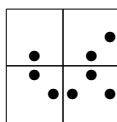


RYSUNEK 13. Niepożądane ustawienia przekaźnika.

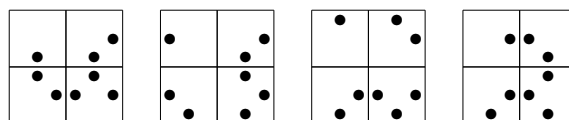
2×2 (czyli rzędu 0) można konstruować pseudobramki wyższych rzędów (również zwane przekaźnikami), w sposób identyczny jak w Obserwacji 3.2.

Nazwa przekaźnika jest zgodna z intuicją — sygnał z lewej krawędzi przechodzi na prawą, a z górnej na dolną.

Pseudobramką jest również poniższy gadżet:

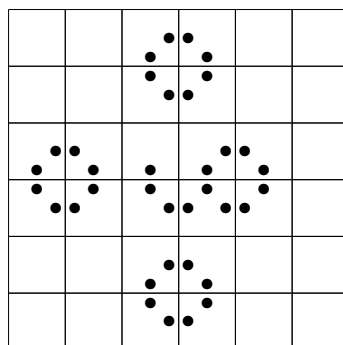


Ma on cztery pokazane na rysunku ustawienia „bramkowe”: 0100, 0111, 1110 i 0010. Podobnie, jak w wypadku przekaźnika, jeśli na dwóch bokach kwadratu wymusimy kolorowanie zgodne z definicją bramki, pozostałe dwie krawędzie również będą prawidłowo (tzn. zgodnie z warunkami narzucenymi na bramkę) pokolorowane. Nazwijmy pseudobramkę o tych czterech stanach **pseudobramką K**.



RYSUNEK 14. Ustawienia „bramkowe” pseudobramki K.

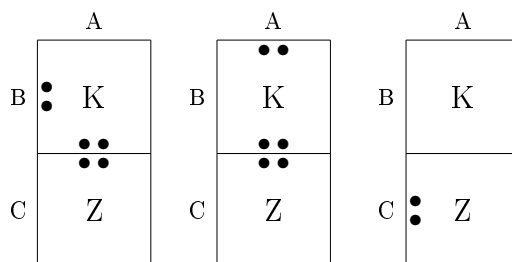
W dalszym ciągu będziemy rozważać jedynie pseudobramki K rozszerzone jak na rysunku 15 (tak, aby prawa krawędź centralnego gadżetu była zawsze aktywna), co eliminuje skrajnie prawe z czterech ustawień zaprezentowanych na rysunku 14.



RYSUNEK 15. Niesymetryczne rozszerzenie pseudobramki K.

Powstały gadżet możemy rozszerzać dalej (symetrycznie, tzn. jak w obserwacji 3.2), tworząc pseudobramki K wyższych rzędów.

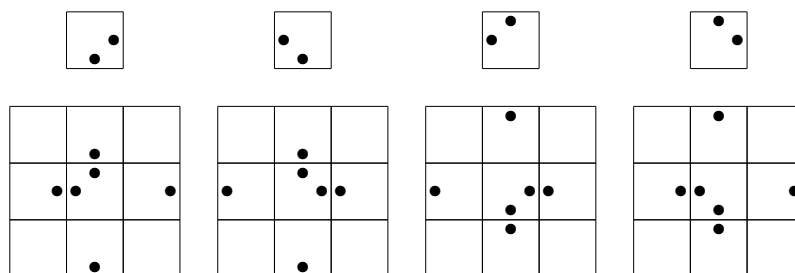
Jeżeli teraz połączymy pseudobramkę K z zakretem negowanym typu lewo-góra (Z), powstały prostokąt będzie zachowywać się jak dwie połączone bramki. Co więcej, z własności gadżetu wynika, że teraz dokładnie jedna z trzech krawędzi A , B , C jest aktywna (rysunek 16). Taką parę nazwiemy **gadżetem klauzuli**, a A , B i C jego **wyróżnionymi krawędziami**.



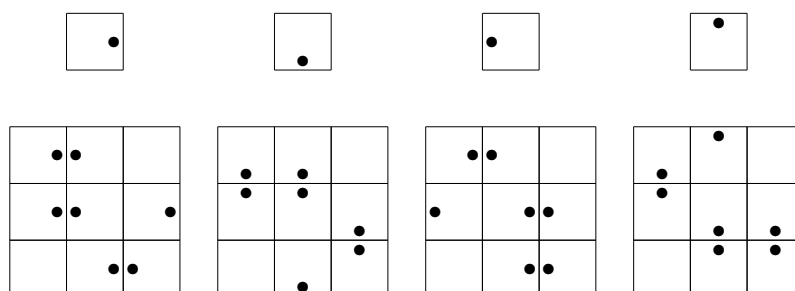
RYSUNEK 16. Gadżet klauzuli

W konstrukcji gadżetu klauzuli użyliśmy płytki T , można ją jednak wyeliminować, jeśli uda nam się skonstruować pseudobramkę K z płytek O , C i D .

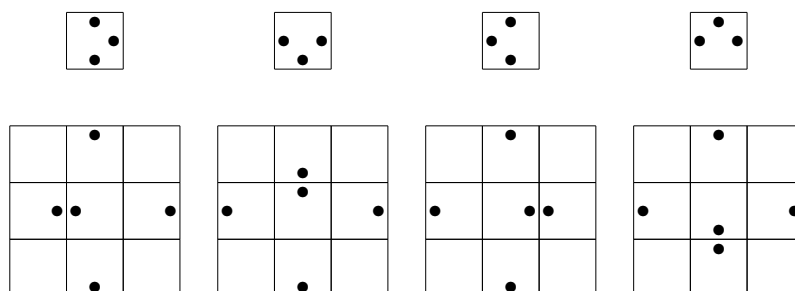
Obserwacja 3.13. *Pseudobramka K jest konstruowalna z płytek O , C i D .*



RYSUNEK 17. Kwadrat \bar{C} – symulacja płytki C .

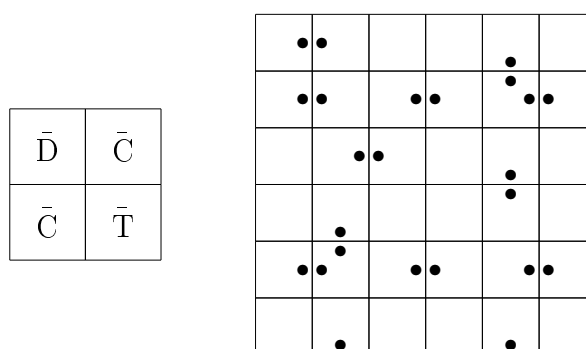


RYSUNEK 18. Kwadrat \bar{D} – symulacja płytki D .

RYSUNEK 19. Kwadrat \bar{T} – symulacja płytki T .

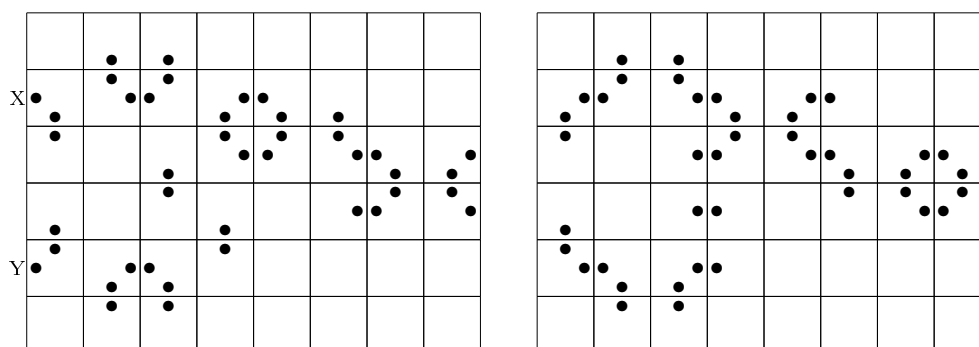
Dowód. Krok 1. Symulujemy zachowanie pojedynczych płytek C , D i T przez kwadraty \bar{C} , \bar{D} i \bar{T} rozmiaru 3×3 , jak na rysunkach 17–19.

Zastępując w pseudobramce K rozmiaru 2×2 płytki C , D i T przez \bar{C} , \bar{D} i \bar{T} (rysunek 20) otrzymujemy kwadrat \bar{K} o boku 6, który od pseudobramki różni się tym, że o stanie krawędzi nie decydują dwa środkowe kwadraty, lecz drugi i piąty z kolei.

RYSUNEK 20. Symulacja pseudobramki K

Krok 2. Do każdej krawędzi kwadratu \bar{K} doklejamy łącznik przedstawiony na rysunku 21. Przenosi on stan środkowych płytek prawej krawędzi na parę druga-piąta lewej.

Tak otrzymaną pseuobramkę K rzędu 5 przedstawia schematycznie rysunek 22. Litera L oznacza łącznik, O - pole wypełnione pustymi płytkami, X, Y — punkty styku kwadratu \bar{K} z łącznikiem. \square



RYSUNEK 21. Łącznik

O	L		O
L	Y	X	L
	\bar{D}	\bar{C}	
O	X	Y	O
	\bar{C}	\bar{D}	
O	L		O

RYSUNEK 22. Pseudobramka K rzędu 5.

3.2. Redukcja.

Twierdzenie 3.14. *Każdy zbiór płytek Z , który:*

- (1) *zawiera czarno-białe płytki O , C i D ,*
- (2) *pozwała skonstruować flip-flop,*

jest układem NP-zupełnym.

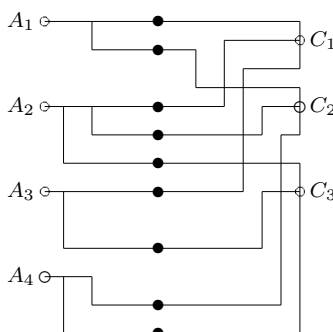
Dowód. Przypomnijmy, że problem (1,3)-SAT (zwany też XSAT od *exact satisfiability*) jest następującym problemem decyzyjnym: mając daną formułę logiczną w postaci 3-CNF rozstrzygamy, czy istnieje wartościowanie

zmiennych takie, by w każdej klauzuli prawdziwy był dokładnie jeden literał. Równoważnie, pytamy o spełnialność formuły postaci

$$\bigwedge_{i=1}^n s(a_{i1}, a_{i2}, a_{i3})$$

gdzie a_{ij} są literałami, zaś s jest (jedynym) trójargumentowym spójnikiem przyjmującym wartość logiczną 1 wtedy i tylko wtedy, gdy dokładnie jeden z jego argumentów przyjmuje wartość 1. Problem (1,3)-SAT jest NP-zupełny [3].

Rozważmy instancję problemu (1,3)-SAT, tzn. formułę $\varphi = \bigwedge_{j=1}^n C_j$, ze zmiennymi x_1, \dots, x_m i klauzulami $C_j = s(a_{j1}, a_{j2}, a_{j3})$, $j = 1, \dots, n$.



RYSUNEK 23. Graf incydencji formuły
 $\varphi = s(x_1, x_2, x_3) \wedge s(x_1, x_2, x_4) \wedge s(x_2, x_3, x_4)$.

Narysujemy teraz na płaszczyźnie tak zwany *graf incydencji* $G(\varphi)$ tej formuły. Składać się on będzie z punktów i fragmentów łamanej, przy czym cały leżeć będzie na kracie \mathbb{Z}^2 (wierzchołki $G(\varphi)$ mają być punktami kratowymi, zaś odcinki zawierać się w prostych kraty).

Krok 1. Na płaszczyźnie zaznaczamy punkty kratowe A_1, \dots, A_m odpowiadające zmiennym, o odciętej równej 0. Dodatkowo dla każdego i powołujemy n punktów A_{i1}, \dots, A_{in} o odciętej m — „kopie” punktu A_i . Prowadzimy następnie dla $i = 1, \dots, m$ połączenie (takie, jak na rysunku 23) od punktu A_i do A_{ij} wtedy i tylko wtedy, gdy zmienna x_i występuje w klauzuli C_j . Jak widać, możemy to zrobić tak, aby nie było skrzyżowań.

Krok 2. Zaznaczamy na płaszczyźnie punkty C_1, \dots, C_n o odciętej $m + 2mn$ odpowiadające klauzulom, a następnie dla każdej klauzuli o numerze j postaci $s(x_{i_1}, x_{i_2}, x_{i_3})$ punkt C_j łączymy z punktami A_{i_1j}, A_{i_2j} i A_{i_3j} . Skrzyżowania łamanych są dozwolone, prowadzimy jednak łamane w taki sposób, żeby żadne dwie nie przecinały się w więcej niż jednym punkcie, na przykład trzema odcinkami – od punktu o współrzędnych (a, b) do punktu (c, d) odcinkami $(a, b) \rightarrow (x, b)$, $(x, b) \rightarrow (x, d)$, $(x, d) \rightarrow (c, d)$, przy czym każdy z pionowych odcinków (będzie ich co najwyżej mn) powinien mieć inną odcięłą.

Krok 3. Niech r będzie rzędem flip-flopa, który jest konstruowalny na mocy założenia o dostępnym zestawie płytek. Z twierdzeń w rozdziale 3.1

wynika, że istnieje takie $k \geq r$, dla którego można z posiadanych płytek oraz flip-flopa skonstruować wszystkie zdefiniowane bramki rzędu k . Cały graf odwzorowujemy na planszy złożonej z płytek, zastępując punkty kratowe i odcinki łamanej przez zespoły bramek rzędu k :

- odcinki i skrzyżowania zastępujemy przez przełączniki
- skrety łamanej przez zakrety
- punkty podziału przez rozdzielniki
- punkty A_1, \dots, A_n przez źródła prawostronne
- punkty B_{ij} przez przełączniki, jeśli są połączone łamaną z innymi punktami, w przeciwnym wypadku — przez puste kwadraty
- punkty C_1, \dots, C_n przez gadżet klauzuli, przy czym bramki odpowiadające trzem odcinkom łamanej wchodzącym do tych punktów powinny sąsiadować z wyróżnionymi krawędziami gadżetu
- pozostałe punkty kraty przez kwadraty pustych płytek

Jeśli zmienna x_i występuje w pewnej klauzuli C_j zanegowana, zastępujemy jeden z zakrętów na łamanej $B_{ij} \rightarrow C_j$ przez zakręt negowany (możemy założyć, że na każdej takiej drodze jest przynajmniej jeden zakręt).

Pokażemy, że cała powstała plansza ma poprawne ustawienie płytek wtedy i tylko wtedy, gdy wyjściowa formuła jest spełnialna.

(\Rightarrow) W poprawnym ułożeniu płytek każde źródło w punkcie A_i jest albo w stanie 0000, albo 0100. W pierwszym przypadku wartościujemy zmienną x_i na 0, w drugim na 1. Z własności bramek wynika, że stan prawej krawędzi źródła A_i jest taki sam, jak lewej i prawej krawędzi przełącznika B_{ij} , dla dowolnego $j = 1, \dots, n$. Teraz, jeśli zmienna x_i występuje w klauzuli C_j , to przełącznik B_{ij} jest połączony za pomocą przełączników z jedną z wyróżnionych krawędzi gadżetu tej klauzuli. Jeśli zmienna ta występuje w klauzuli zanegowana, to na drodze jeden z zakrętów jest negowany, a zatem stan tej krawędzi jest przeciwny do stanu B_{ij} , a więc i do stanu źródła A_i . Jeśli zaś zanegowana nie jest, to stan tej wyróżnionej krawędzi musi być zgodny ze stanem A_i . Wiemy jednak, że dokładnie jedna z trzech krawędzi jest w stanie 1. A zatem przy podanym przypisaniu wartości zmiennych dokładnie jeden z literałów klauzuli C_j jest prawdziwy. Każda klauzula jest więc przy takim wartościowaniu zmiennych prawdziwa, a zatem formuła jest spełnialna.

(\Leftarrow) Jeśli formuła jest spełnialna, to orientujemy źródło A_i tak, aby było zgodne z wartością logiczną zmiennej x_i w wartościowaniu w , które czyni formułę prawdziwą: prawa krawędź źródła A_i jest aktywna wtedy i tylko wtedy, gdy $w(x_i) = 1$. Wszystkie pozostałe bramki na planszy, z wyjątkiem być może gadżetów klauzul mają teraz dokładnie jedną możliwą orientację (przełączniki, zakrety i rozdzielniki zastępujące fragmenty łamanej mają narzucone stany przynajmniej dwóch krawędzi, a to narzuca im ustawienie). Podobnie jak w dowodzie przeciwnej implikacji dowodzimy, że odpowiednie krawędzie w gadżetach klauzul są aktywne wtedy i tylko wtedy, jeśli odpowiadające im bramki zmiennych są aktywne (lub nieaktywne, jeśli zmienna występuje w klauzuli zanegowana). Ale skoro formuła jest

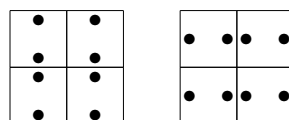
prawdziwa przy wartościowaniu w , to dokładnie jedna krawędź będzie aktywna, a to oznacza, że potrafimy obie bramki gadżetu klauzuli zorientować poprawnie. Wszystkie bramki są teraz w poprawnym ustawieniu.

Graf incydencji wymaga $O((mn)^2)$ bramek, a każda bramka to k^2 płytek (k zależy od posiadanego zestawu płytek, jest więc stałą). Transformacja (1,3)-SAT do problemu hydraulika jest zatem wielomianowa, co dowodzi NP-zupełności tego ostatniego. \square

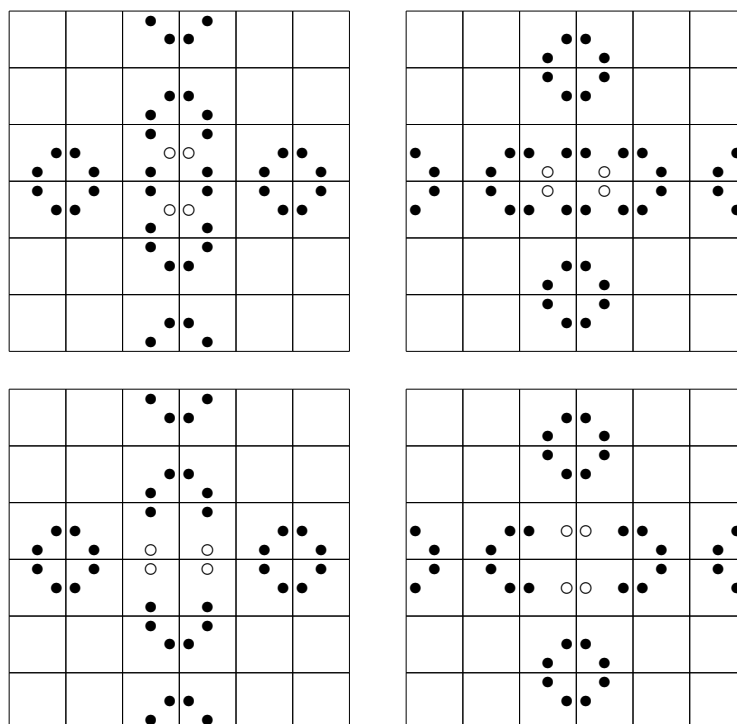
3.3. Wnioski.

Wniosek 3.15. *Zestaw płytek OCDS jest NP-zupełny.*

Dowód. Kwadrat 2×2 złożony z płytek S jest bramką flip-flop rzędu 0. \square

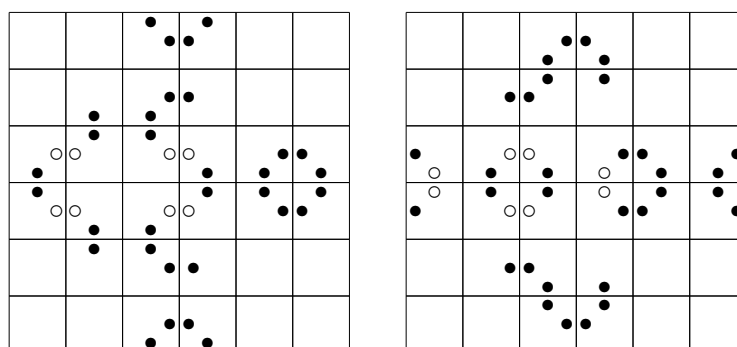


RYSUNEK 24. Flip-flop zbudowany z płytek typu S .



RYSUNEK 25. Flip-flop zbudowany z płytek typu O, C, D i 1210 (górne dwa rysunki) oraz płytek typu O, C, D i 0201 (rysunki dolne).

Wniosek 3.16. *Zestaw zawierający płytki O, C, D i jedną z płytek: 1210 lub 0201 jest NP-zupełny.*



RYSUNEK 26. Flip-flop zbudowany z płytek typu O , C , D , 1002 i 2001.

Dowód. Odpowiednie bramki flip-flop przedstawia rysunek 25. □

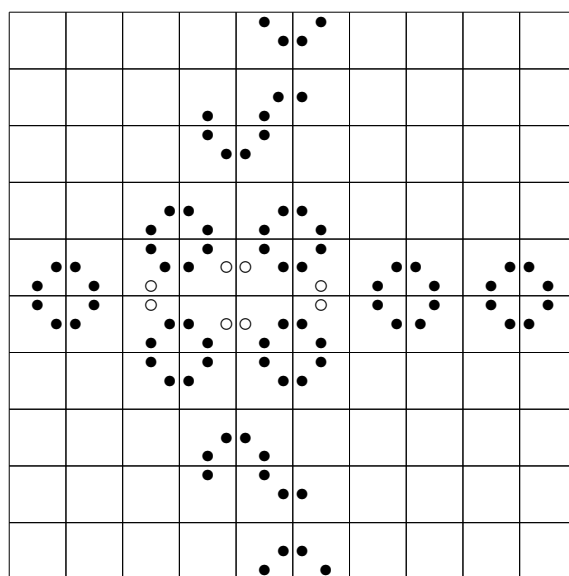
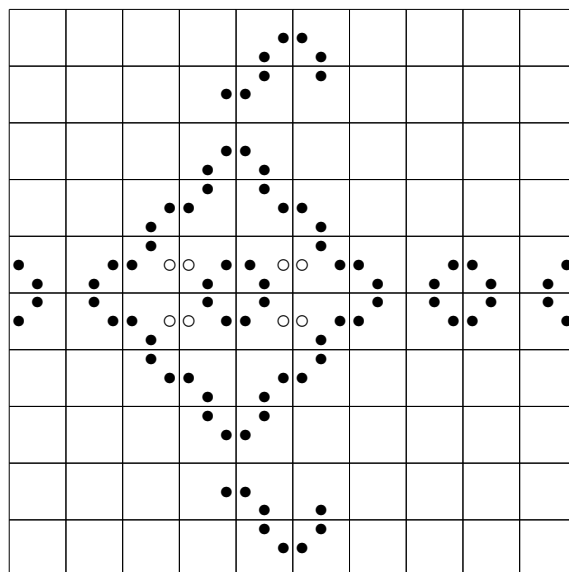
Wniosek 3.17. Zestaw zawierający płytki O , C , D oraz 1002 i 2001 jest NP -zupełny.

Dowód. Odpowiednią bramkę flip-flop przedstawia rysunek 26. □

Wniosek 3.18. Zestaw zawierający płytki O , C , D oraz 0112 i 2110 jest NP -zupełny.

Dowód. Odpowiednią bramkę flip-flop (rzędu 2) przedstawia rysunek 27. □

Wystarczy zatem do wielomianowego zestawu płytek OCD dołączyć, symetryczną (mającą oś symetrii) trójkolorową płytkę, lub parę lustrzanych płytek niesymetrycznych, aby problem hydraulika zrobił się NP -zupełny. Jak przekonamy się w twierdzeniu 4.2, dołączenie tylko jednej płytki niesymetrycznej pozostawia problem w P .



RYSUNEK 27. Flip-flop zbudowany z płytek typu O , C , D , 0112 i 2110 .

4. Trzy kolory — przypadki wielomianowe

Twierdzenie 4.1. *Zestaw płytek postaci $\mathcal{BW} \cup \mathcal{BR} \cup \mathcal{WR}$, gdzie \mathcal{BW} zawiera tylko płytki czarno-białe, \mathcal{BR} — czarno-czerwone, a \mathcal{WR} — biało-czerwone, jest wielomianowy wtedy i tylko wtedy, gdy każdy z tych trzech zestawów jest wielomianowy.*

Dowód. Konieczność warunku jest oczywista. Pokażemy jego dostateczność: rozważmy dowolny spójny obszar \mathcal{O} na planszy, który zawiera tylko płytki z zestawu \mathcal{BW} i jest maksymalny, tzn. wszystkie płytki sąsiadujące z płytkami z \mathcal{O} należą do \mathcal{BR} lub \mathcal{WR} . Zauważmy, że wszystkie krawędzie tego obszaru sąsiadujące z płytkami z \mathcal{BR} muszą być czarne, a wszystkie sąsiadujące z \mathcal{WR} - białe. Jest to więc problem dwukolorowego hydraulika z ustalonym brzegiem, na mocy twierdzenia 2.5 dający się rozwiązać w czasie wielomianowym. Jeśli odpowiedź jest negatywna (płytek nie da się prawidłowo ułożyć), oczywiście cała plansza również nie daje się poprawnie ułożyć. Podobnie pokazujemy, że każdy maksymalny spójny obszar płytek z \mathcal{BR} lub \mathcal{WR} to w istocie instancja problemu dwukolorowego hydraulika z ustalonym brzegiem. Pełny algorytm wygląda zatem następująco:

1. Dzielimy planszę na maksymalne spójne obszary płytek z jednego zestawu (np. za pomocą algorytmu DFS na grafie, którego wierzchołkami są pola, a krawędzie poprowadzone są między płytkami z tego samego zestawu)
2. Dla każdego obszaru ustalamy konieczne stany płytek brzegowych i rozstrzygamy odpowiadający mu problem dwukolorowego hydraulika.
3. Zwracamy „TAK” jeśli w kroku 2 na wszystkich obszarach odpowiedź brzmiała „TAK”, w przeciwnym przypadku zwracamy „NIE”.

□

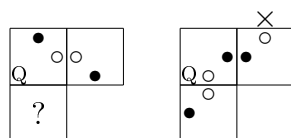
Twierdzenie 4.2. *Zestaw płytek $OCDXTA$, gdzie A jest jedną z płytek 1002, 2001, 0112 lub 2110 jest wielomianowy.*

Dowód. Bez straty ogólności przyjmijmy, że A jest płytką 1200, dla innych przypadków dowód jest analogiczny.

Pokażemy, że każda z płytek A może przyjmować tylko jedną możliwą orientację: nie istnieją dwa poprawne ustawienia planszy, przy których któraś z trój kolorowych płytek mogłaby się znajdować w dwóch różnych położeniach.

Założmy niewprost, że to nieprawda — istnieją płytki A o dwóch możliwych położeniach. Wybierzmy więc najwyższy wiersz, w którym są takie płytki, a w nim płytkę najbardziej z lewej i oznaczmy ją przez Q . Płytką na lewo od Q jest albo czarno-biała, albo jest płytką A o jednej możliwej orientacji, a wtedy jej czerwona krawędź nie może być sąsiednia do Q (narzucałaby jedyną możliwą orientację dla Q). Podobnie płytką bezpośrednio nad Q musi mieć sąsiadującą z nią krawędź czarną lub białą. To pozostawia płytce Q dwa ustawienia, pokazane na rysunku.

Aby oba były możliwe, pozostałe dwie sąsiednie płytki muszą być typu A , a to z kolei oznacza, że w jednym (i tylko jednym) z tych ustawień płytką na prawo od Q musi mieć górną krawędź czerwoną. W wyższym wierszu



RYSUNEK 28. Płytki Q ma tylko dwa możliwe ustawienia; drugie prowadzi do sprzeczności.

musi więc znajdować się płytki typu A o dwóch możliwych orientacjach — sprzeczność z wyborem płytki Q .

Z tego dowodu wynika, że możemy dla każdej płytki A wyznaczyć jej jedyne możliwe położenie — zaczynając od najwyższych rzędów i przeglądając płytki od lewej. Skoro wszystkie płytki A mają ustalone położenia, możemy je pogrupować parami tak, jak stykają się przez swoje czerwone krawędzie, po czym zastąpić, również ustalonymi, płytkami czarno-białymi (kolor czerwony zastępujemy białym). Pozostaje rozwiązać dwukolorowy problem hydraulika z ustalonym brzegiem, co jesteśmy w stanie zrobić w czasie wielomianowym (Twierdzenie 2.5). \square

5. Podsumowanie

Problem dwukolorowego hydraulika jest rostrzygnięty dla wszystkich zestawów płytek \mathcal{Z} , z wyjątkiem przypadku, gdy $\{D, S\} \subseteq \mathcal{Z} \subseteq \{D, S, O, T, X\}$. Zestawy zawierające płytki C , S i D są NP-zupełne, zaś inne — wielomianowe [1]. Modyfikacja dowodu z [1] pozwala uogólnić go na niektóre istotne przypadki tych zestawów trój kolorowych, które zawierają płytki O , C i D . Pokazaliśmy też, że nie istnieją zestawy NP-zupełne bez płytek istotnie trój kolorowych.

Cały czas otwartym problemem pozostaje złożoność obliczeniowa problemu hydraulika dla zestawu DS , jak również pełna charakteryzacja trój kolorowych zestawów NP-zupełnych.

Podziękowania

Składam serdeczne podziękowania Panu Profesorowi dr hab. Pawłowi Idziakowi za dokładne przedyskutowanie tematu, opiekę nad pracą i pomoc w jej zredagowaniu. Dziękuję również kolegom za wspaniałą atmosferę seminariów Katedry Algorytmiki oraz wiele cennych dyskusji – szczególnie dziękuję Kamilowi Klochowi oraz Bartkowi Walczakowi.

Literatura

- [1] D. Kral, V. Majerech, J. Sgall, T. Tichy and G. Woeginger, It is tough to be a plumber, *Theoretical Computer Science* **313**(2004) 473–484.
- [2] J. Hopcroft and R. Karp, An $n^{5/2}$ algorithm for matching in bipartite graphs, *SIAM, J. Comput.* **2**(4) (1975) 225–231.
- [3] J. Schaefer, The complexity of satisfiability problems *Proc. 10th Ann. ACM Symp. on Theory of Computing, 1978*, Association for Computing Machinery, New York, 216–226.