

O dziwnym dodawaniu

Leszek PIENIAŹEK, Jacek TABOR, Kraków

1. Motywacja

Chcielibyśmy zacząć referat od przedstawienia z życia wziętej motywacji problemu który będziemy badać:

Problem 1. Niech dany będzie krasnoludek Gomlin... Tak się często zaczyna każdy tekst matematyczny, więc i ja tak zacząłem. Kontynuujmy więc.

Gomlin mieszka na brzegu rzeki (szerokość rzeki wynosi l), i chce się przeprawić na drugi brzeg (celem wyjaśnienia: tam mieszka krasnoludka Galia). Musi więc zbudować most (oczywiście o długości przynajmniej l), a dysponuje prętami metalowymi o długościach l_1, \dots, l_n .

Niestety, przy łączeniu dwóch metalowych prętów o długościach d_1, d_2 , kowal pobiera opłatę odcinając 10% długości powstałego pręta.

Pytania które stawia sobie Gomlin są następujące:

- czy zgromadzone pręty wystarczą? czy też trzeba dalej pracować i kupować nowe?
- w jakiej kolejności łączyć ze sobą te pręty, aby otrzymać most możliwie największej długości?

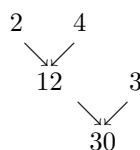
Drugi przykład który będzie prowadził nas do podobnych rozważań jest następujący:

Problem 2. Załóżmy, że chcemy zbudować drzewo mając dane liście o wagach w_1, \dots, w_n . Możemy utworzyć nową gałązkę łącząc ze sobą dowolne dwie gałązki, ale waga łączenia jest równa sumie wag łączonych gałązek (inaczej mówiąc waga powstałej gałązki jest równa dwukrotności wag łączonych gałązek).

Dla przykładu, z liści o wadze 1 i 2 utworzymy następujące drzewo:



(jak widzimy waga powstałego drzewa wynosi 6), z liści o wadze 2, 3, 4 możemy utworzyć między innymi takie drzewo:



(najpierw łączymy ze sobą liście o wagach 2 i 4 tworząc gałązkę o wadze 12, a następnie łączymy powstałą gałązkę z liściem o wadze 3).

I pytanie teraz, w jaki sposób tworzyć nasze drzewo, aby jego waga była możliwie najmniejsza?

Warto tutaj jeszcze nadmienić, że tego typu problemy pojawiają się także w topologii ogólnej przy okazji tak zwanych quasi-norm [5, 1] i quasi-metryk [2, 3]. Opowiemy więc jeszcze troszkę o quasi-metrykach.

Zacznijmy od przypomnienia pojęcia metryki. Otóż *metryka* jest to funkcja d , która mierzy nam odległość między dwoma punktami. Spełnia ona (między innymi) tak zwaną *nierówność trójkąta*, to znaczy

$$d(A, C) \leq d(A, B) + d(B, C)$$

dla dowolnych punktów A, B, C . Jeśli teraz mamy punkty A_1, \dots, A_n , to oczywiście

$$d(A_1, A_n) \leq l_1 + \dots + l_{n-1},$$

gdzie $l_i = d(A_i, A_{i+1})$.

Wyobraźmy sobie teraz, że zamiast mierzenia odległości między dwoma punktami będziemy mierzyć czas potrzebny nam na przejście tej odległości. Czy wtedy ten czas będzie spełniał nierówność trójkąta? Oczywiście, że nie! – gdyż się po prostu troszkę męczymy (jeśli przejdziemy 5 km w godzinę, to wcale nie znaczy, że przejdziemy 10 km w dwie godziny). Musimy więc zmodyfikować warunek trójkąta. Najprostszy sposób, w jaki możemy to zrobić, to założyć, że nasza funkcja d spełnia warunek

$$d(A, C) \leq K(d(A, B) + d(B, C)) \text{ dla dowolnych } A, B, C,$$

gdzie $K > 1$ jest ustalone (i zależy od naszej kondycji fizycznej – im K jest bliższe jeden, tym, oczywiście, nasza kondycja jest lepsza).

Teraz widzimy już naturalny problem dotyczący quasi-metryk:

Problem 3. Mamy dane punkty A_1, \dots, A_n . Niech $l_i = d(A_i, A_{i+1})$, gdzie d jest quasi-metryką. Jakie jest (możliwie) najlepsze oszacowanie na

$$d(A_1, A_n)?$$

2. Dziwne dodawanie

Jeśli dobrze się przypatrzeć, to wszystkie wyżej opisane problemy prowadzą nas do następującej definicji *dziwnego dodawania*:

Definicja 1. Niech $K > 0$ będzie dane. Dla $x, y \in \mathbb{R}_+$ definiujemy

$$x \oplus_K y = K \cdot (x + y).$$

Oczywiście, w przypadku gdy $K = 1$ dostajemy zwykłe dodawanie. Przypatrzmy się teraz, jak możemy wyabstrahować problemy opisane w poprzednim rozdziale za pomocą naszego dziwnego dodawania.

Krasnoludek: Mamy dane liczby l_1, \dots, l_n , $K = 0,9$. Chcemy je tak dodać (w dowolnej kolejności) za pomocą dodawania \oplus_K , aby uzyskać możliwie największą liczbę (i najlepiej dla przetrwania gatunku krasnoludków, aby była ona większa niż l).

Drzewo: Mamy dane wagi w_1, \dots, w_n , $K = 2$. Mamy je tak dodać (w dowolnej kolejności) za pomocą dodawania \oplus_K , aby uzyskać możliwie najmniejszą wagę.

Quasi-metryka: Dane są liczby l_1, \dots, l_{n-1} , $K > 1$. Mamy je wszystkie dodać tak, aby utworzyć możliwie najmniejszą sumę, ale wolno nam dodawać tylko w tej kolejności, która jest (czyli możemy do l_2 dodać l_3 , ale nie możemy do l_1 dodać l_3).

Widzimy, że powstają nam tak naprawdę dwa problemy – problem przemienny, gdzie możemy dodawać każdą liczbę do każdej, i nieprzemienny, gdzie nie możemy zmienić kolejności. Zobaczmy może, jak to działa na przykładach. Dla prostoty przyjmujemy, że $K = 2$.

Przykład 1 (przemienny). Chcemy dodać 5, 3 i 1. Możemy to zrobić tak: $5 \oplus_2 3 = 2 \cdot (5 + 3) = 16$, i potem $16 \oplus_2 1 = 34$. Ale łatwo zauważyć, że to nie jest najlepszy wynik – lepiej zrobić tak:

$$(1 \oplus_2 3) \oplus_2 5 = 8 \oplus_2 5 = 26.$$

Łatwo można sprawdzić, że 26 jest najlepszym możliwym wynikiem, który można otrzymać z przemiennego dodawania 1, 3 i 5 za pomocą operacji \oplus_2 .

Jako proste zadanie zostawiamy dla Czytelnika znalezienie najmniejszej możliwej dziwnej sumy dla liczb 1, 2, 3, 4.

Przykład 2 (nieprzemienny). Spróbujmy teraz dodać (w sposób nieprzemienny) liczby 4, 3, 2, 1. Łatwo zauważyć, że to co musimy wykonać, to takie „wrzucenie” nawiasów do ciągu $4 \oplus_2 3 \oplus_2 2 \oplus_2 1$, aby było „rozsądnie”. Można to zrobić na parę możliwych sposobów (pojawiają się tutaj tak zwane liczby Catalana).

Popatrzmy więc (jeśli się nie pomyliliśmy, to poniżej są wypisane wszelkie możliwe wstawienia nawiasów):

$$\begin{aligned} \{(4 \oplus_2 3) \oplus_2 2\} \oplus_2 1 &= \{14 \oplus_2 2\} \oplus_2 1 = 32 \oplus_2 1 = 66, \\ \{4 \oplus_2 (3 \oplus_2 2)\} \oplus_2 1 &= ? \\ (4 \oplus_2 3) \oplus_2 (2 \oplus_2 1) &= ? \\ 4 \oplus_2 \{(3 \oplus_2 2) \oplus_2 1\} &= ? \\ 4 \oplus_2 \{3 \oplus_2 (2 \oplus_2 1)\} &= ? \end{aligned}$$

Wykonaliśmy pierwsze działanie, i wyszło 66. Zostawiamy dla zainteresowanego Czytelnika wykonanie pozostałych zadań i znalezienie tej, w której suma jest najmniejsza.

3. Troszkę paskudnej matematyki

Od tego momentu będziemy zakładali, że $K > 1$ (sytuacja, gdy $K < 1$, jest analogiczna).

Dla ciągu $x = (x_1, \dots, x_n)$ przez $\mathcal{P}x$ będę oznaczał najmniejszą wartość możliwą do uzyskania za pomocą dodawania przemiennej, a przez $\mathcal{N}x$ najmniejszą wartość możliwą do uzyskania z dodawania przemiennej.

Zajmiemy dwoma głównymi pytaniami:

- jaki jest algorytm obliczania $\mathcal{P}x$ oraz $\mathcal{N}x$?
- jakie są oszacowania wartości $\mathcal{P}x$ oraz $\mathcal{N}x$?

W tym rozdziale zajmiemy się próbą oszacowania (z góry i z dołu) wartości $\mathcal{P}x$ i $\mathcal{N}x$.

Przydatna będzie następująca obserwacja:

Stwierdzenie 1. *Weźmy liczby rzeczywiste nieujemne x_1, \dots, x_n . Wtedy wszystkie wyniki, możliwe do uzyskania z dodawania przemiennej, dane są przez zbiór*

$$\mathcal{S} := \{K^{l_1}x_1 + \dots + K^{l_n}x_n \mid l_i \in \mathbb{Z}, 1/2^{l_1} + \dots + 1/2^{l_n} = 1\}.$$

Dowód. Przeprowadzimy dowód indukcyjny ze względu na długość ciągu. Oczywiście twierdzenie zachodzi dla $n = 1$.

Założmy, że twierdzenie jest prawdziwe dla pewnego n . Pokażemy, że jest prawdziwe dla $n + 1$.

De facto mamy do pokazania dwie rzeczy: że każdy wynik dziwnego dodawania jest realizowany przez element zbioru \mathcal{S} , i odwrotnie, że każdy element z \mathcal{S} daje się zrealizować jako wynik dziwnego dodawania.

Pokażę najpierw, że każdy element dziwnego dodawania należy do \mathcal{S} . Weźmy więc x_1, \dots, x_{n+1} i założmy, że przez dziwne dodawanie tego ciągu możemy uzyskać $r \in \mathbb{R}$. Wtedy oczywiście biorąc pod uwagę ostatnią operację otrzymujemy, że

$$r = s \oplus_K t,$$

gdzie s jest dziwną sumą $x_{\sigma(1)}, \dots, x_{\sigma(k)}$, a t jest dziwną sumą $x_{\sigma(k+1)}, \dots, x_{\sigma(n+1)}$ dla pewnej permutacji σ zbioru $\{1, \dots, n+1\}$ oraz $k \in \{1, \dots, n\}$. Na podstawie założenia indukcyjnego

$$s = K^{l_1}x_{\sigma(1)} + \dots + K^{l_k}x_{\sigma(k)}, \quad t = K^{l_{k+1}}x_{\sigma(k+1)} + \dots + K^{l_{n+1}}x_{\sigma(n+1)},$$

dla pewnego ciągu liczb naturalnych l_1, \dots, l_{n+1} takiego, że

$$(1) \quad 1/2^{l_1} + \dots + 1/2^{l_k} = 1, \quad 1/2^{l_{k+1}} + \dots + 1/2^{l_{n+1}} = 1.$$

W konsekwencji widzimy, że

$$r = K(s + t) = K^{l_1+1}x_{\sigma(1)} + \dots + K^{l_{n+1}+1}x_{\sigma(k)},$$

a na podstawie (1) otrzymujemy, że $1/2^{l_1+1} + \dots + 1/2^{l_{n+1}+1} = 1$.

Tak naprawdę, to skupimy się tylko na sytuacji przemiennej, ponieważ jest dużo prostsza.

Dowód faktu, że każdy element ze zbioru \mathcal{S} daje się uzyskać za pomocą pewnej dziwnej sumy pozostawiam do pokazania Czytelnikowi. Jako wskazówkę jedynie dodam, że opiera się on na następującej prostej, ale sympatycznej obserwacji:

Sympatyczna obserwacja. Jeśli liczby naturalne l_1, \dots, l_{n+1} spełniają warunek

$$1/2^{l_1} + \dots + 1/2^{l_{n+1}} = 1,$$

to istnieje taka permutacja σ zbioru $\{1, \dots, n+1\}$ oraz takie $k \in \{1, \dots, n\}$, że

$$1/2^{l_{\sigma(1)}} + \dots + 1/2^{l_{\sigma(k)}} = 1/2, \quad 1/2^{l_{\sigma(k+1)}} + \dots + 1/2^{l_{\sigma(n+1)}} = 1/2.$$

□

Do jednego z naszych głównych wyników przyda się nam następująca konwencja: dla $w \in (0, 1)$ i ciągu $x = (x_1, \dots, x_n)$ kładziemy

$$\|x\|_w := (x_1^w + \dots + x_n^w)^{1/w}.$$

Dygresja 1. I tu, chcąc, nie chcąc, wypadły nam z rękawa quasi-normy. Dla kogoś wiedzącego, co to jest norma, chciałbym nadmienić, że $\|\cdot\|_w$, normą nie jest (choć tak na pierwszy rzut oka może się wydawać), gdyż nie spełnia warunku trójkąta – mamy jedynie

$$\|x + y\|_w \leq 2^{1/w-1}(\|x\|_w + \|y\|_w).$$

Powód, dla którego $\|\cdot\|_w$ nie jest normą, łatwo zauważyć rysując na płaszczyźnie koło jednostkowe (okazuje się, że nie jest ono wypukłe).

I teraz możemy już sformułować

Twierdzenko 1. Niech będzie dany ciąg $x = (x_1, \dots, x_n)$. Niech $w \in (0, 1)$ będzie takie, że

$$(2) \quad 1 \oplus_K 1 = \|(1, 1)\|_w.$$

Wtedy

$$\|x\|_w \leq \mathcal{P}x \leq K\|x\|_w,$$

i powyższych oszacowań nie da się poprawić.

Zanim przejdę do dowodu, chciałbym nadmienić, że w spełniające (2) wyraża się jawnym wzorem $w = 1/\log_2(2K)$, i ponieważ $K > 1$, więc otrzymujemy oczywiście, że $w \in (0, 1)$.

Szkic dowodu. Zaczniemy, od pokazania, że nasze szacowanie jest najlepsze z możliwych. Mamy

$$\mathcal{P}(1, 1) = 1 \oplus_K 1 = \|(1, 1)\|_w \quad \text{oraz} \quad \mathcal{P}(0, 1) = K = K \cdot \|(0, 1)\|_w.$$

A teraz spróbujemy pokazać ideę tego, że ono zachodzi dla dowolnych ciągów. Na podstawie Stwierdzenia 1 mamy

$$\mathcal{P}x = \inf\{K^{l_1}x_1 + \dots + K^{l_n}x_n \mid l_i \in \mathbb{Z}, 1/2^{l_1} + \dots + 1/2^{l_n} = 1\}.$$

W konsekwencji oczywiście

$$\mathcal{P}x \geq \inf\{K^{w_1}x_1 + \dots + K^{w_n}x_n \mid w_i \in \mathbb{R}, 1/2^{w_1} + \dots + 1/2^{w_n} = 1\}.$$

Korzystając z metody szukania minimum funkcji odkrytej przez Lagrange'a (tak zwane mnożniki Lagrange'a) można łatwo pokazać, że prawa strona powyższej nierówności jest realizowana dla pewnych $w_1, \dots, w_n \in (0, \infty)$ i wynosi dokładnie $\|x\|_w$. Uff... A więc mamy

$$\mathcal{P}x \geq \|x\|_w.$$

Pozostało nam jeszcze do pokazania, że $\mathcal{P}x \leq K\|x\|_w$. Ale to także będzie wynikało z powyższego wzoru. Wiemy więc, że istnieją liczby $w_1, \dots, w_n \in (0, \infty)$ takie, że

$$1/2^{w_1} + \dots + 1/2^{w_n} = 1,$$

oraz, że

$$\|x\|_w = K^{w_1}x_1 + \dots + K^{w_n}x_n.$$

Gdyby przez przypadek okazało się, że liczby w_1, \dots, w_n są naturalne, to mielibyśmy koniec dowodu (niestety, życie takie piękne nie jest).

Co więc możemy zrobić? „Znaturalmmy” te liczby i będzie OK. Kładziemy więc

$$l_1 := \lceil w_1 \rceil, \dots, l_n := \lceil w_n \rceil,$$

gdzie przez $\lceil w \rceil$ rozumiemy najmniejszą liczbę całkowitą większą bądź równą w . Wtedy mamy

$$K\|x\|_w = K \cdot (K^{w_1}x_1 + \dots + K^{w_n}x_n) \geq K^{l_1}x_1 + \dots + K^{l_n}x_n.$$

Gdybyśmy mieli (co zazwyczaj nie zachodzi), że $1/2^{l_1} + \dots + 1/2^{l_n} = 1$, to na podstawie Stwierdzenia 1 byłby koniec, gdyż $\mathcal{P}x$ byłoby mniejsze niż $K^{l_1}x_1 + \dots + K^{l_n}x_n$.

Na szczęście dla nas

$$1/2^{l_1} + \dots + 1/2^{l_n} < 1.$$

Pozwala to w sposób sztuczny „dorzucić” do ciągu l_1, \dots, l_n pewną ilość elementów l_{n+1}, \dots, l_m ze zbioru liczb naturalnych, tak aby

$$1/2^{l_1} + \dots + 1/2^{l_m} = 1.$$

Dalej tworzymy m -elementowy ciąg y za pomocą formuły

$$y := (x_1, \dots, x_n, 0, \dots, 0).$$

Otrzymujemy wtedy, na podstawie wcześniejszych obserwacji, że

$$\mathcal{P}y \leq K^{l_1}y_1 + \dots + K^{l_m}y_m = K^{l_1}x_1 + \dots + K^{l_n}x_n \leq K\|x\|_w.$$

Jedyne, co należy jeszcze zauważyć (i co wcale, wbrew pozorom, nie jest trywialne), to fakt, że z konstrukcji y wynika

$$\mathcal{P}x \leq \mathcal{P}y.$$

I koniec. □

Można także pokazać, chociaż jest to trudniejsze, oszacowania dla sytuacji nieprzemiennej:

Twierdzenko 2. Niech będzie dany ciąg $x = (x_1, \dots, x_n)$. Niech $w \in (0, 1)$ będzie takie, że

$$1 \oplus_K 1 = \|(1, 1)\|_w.$$

Wtedy

$$\|x\|_w \leq \mathcal{P}x \leq K^2\|x\|_w,$$

i powyższego oszacowania nie da się poprawić.

4. Troszkę ładnej matematyki

Mamy zatem oszacowania wartości $\mathcal{P}x$ (i $\mathcal{N}x$). Wracamy więc do naszego podstawowego pytania:

Czy krasnoludek jest już szczęśliwy?

Otóż niestety nie. Na podstawie naszych twierdzeń możemy już powiedzieć Krasnoludkowi: tak, tyle prętów już wystarcza do zbudowania mostu, ale niestety, nasza odpowiedź jest, jak to niestety często w stosowaniu matematyki bywa, całkowicie nieprzydatna do krasnoludka.

Mianowicie ciągle nie wiemy, w jakiej kolejności należy spawać te pręty, aby zbudować ten most.

Powstaje więc problem algorytmu, jak zbudować optymalny most.

Algorytm dla sytuacji przemiennej. Otóż okazuje się, że krasnoludek powinien spawać u kowala zawsze dwa najkrótsze pręty spośród tych, które posiada.

Sytuacja gdy $K > 1$ (przeciwna do krasnoludka, dla którego $K = 0,9$), jest analogiczna. Aby otrzymać najmniejszą możliwą sumę należy zawsze dodawać najmniejsze liczby.

Przykład 3. Weźmy liczby 1, 3, 4, 5, 6, 10 i dodajmy je za pomocą powyższego algorytmu dla dodawania \oplus_2 tak, aby uzyskać najmniejszą sumę:

$$\begin{aligned} 1 \oplus_2 3 &= 8, & 4, 5, 6, 10; \\ 4 \oplus_2 5 &= 18, & 6, 8, 10; \\ 6 \oplus_2 8 &= 28, & 10, 18; \\ 10 \oplus_2 18 &= 56, & 28; \\ 28 \oplus_2 56 &= 168. \end{aligned}$$

Porównajmy ten wynik jeszcze z oszacowaniem z poprzedniego rozdziału. Łatwo wyliczamy, że $w = 1/2$, i konsekwencji mamy

$$\|(1, 3, 4, 5, 6, 10)\|_{1/2} \leq \mathcal{P}(1, 3, 4, 5, 6, 10) \leq 2\|(1, 3, 4, 5, 6, 10)\|_{1/2},$$

a ponieważ

$$\|(1, 3, 4, 5, 6, 10)\|_{1/2} = (\sqrt{1} + \sqrt{3} + \sqrt{4} + \sqrt{5} + \sqrt{6} + \sqrt{10})^2 \cong 158,2535366,$$

więc otrzymujemy, że $\mathcal{P}(1, 3, 4, 5, 6, 10)$ jest dosyć bliskie $\|(1, 3, 4, 5, 6, 10)\|_{1/2}$. Okazuje się, że tak się będzie dość często zdarzać (kiedy elementy w danym ciągu są mniej więcej tego samego rzędu).

Na uzasadnienie, dlaczego ten algorytm jest dobry, naprowadza nas Problem 2. Otóż wygodnie nam będzie pracować na drzewach. Z każdym dziwnym dodawaniem liczb w_1, \dots, w_n możemy związać drzewo (binarne) T , takie że z liśćmi drzewa utożsamiamy wszystkie elementy ciągu w_i .

Ustalmy więc teraz takie drzewo T . Dla dowolnej „gałązki” G (a formalnie poddrzewa) z drzewa T definiujemy (indukcyjnie) jej wagę $w(G)$ jako sumę wag tych dwóch gałązek, które po połączeniu tworzą naszą gałązkę przemnożoną przez K (waga i -tego liścia wynosi oczywiście w_i). Nas będzie oczywiście interesowała waga całego drzewa.

Okazuje się teraz, że dowód faktu, iż nasz algorytm jest dobry (którego to dowodu z braku miejsca nie zdołamy przedstawić) opiera się na dwóch podstawowych faktach znanych dobrze każdemu ogrodnikowi:

- wymiana ze sobą (przeszczepianie) tych gałązek w drzewie, które są na tej samej wysokości, nie zmienia wagi drzewa;
- jeśli wymienimy gałązkę G , która jest zaczepiona wyżej, z gałązką H , która jest zaczepiona niżej, to waga drzewa się zmniejszy, jeśli $w(G) > w(H)$.

Na koniec chcielibyśmy jeszcze wspomnieć, że żaden nietrywialny algorytm pozwalający na uzyskanie minimalnej sumy w sytuacji nieprzemiennej nie jest nam znany.

Literatura

- [1] Y. Benyamini, J. Lindenstrauss, *Geometric Nonlinear Functional Analysis, Vol. 1*, AMS Colloquium Publications 48.
- [2] P. Fletcher, W. Lindgren, *Quasi-Uniform Spaces*, Dekker, New York 1982.
- [3] R. Macias, C. Segovia, *Lipschitz Functions on Spaces of Homogenous Type*, *Advances in Mathematics* **33**, 257–270 (1979).
- [4] L. Pieniążek, J. Tabor, *On strange addition*, preprint.
- [5] S. Rolewicz, *Metric linear spaces*, Polish Scientific Publishers and D. Reidel, Dordrecht, 1984.

Precyzyjniej mówiąc, będziemy pracować na pewnym rodzaju grafów zwanych drzewami binarnymi.

Nietrywialny, czyli taki, który nie sprawdza wszystkich możliwości.