Złoty podział w sortowaniu

Marcin PECZARSKI[*]

Złoty podział odcinka, zwany też złotą proporcją, jest doskonale znany. Okazuje się, że podobną własność można sformułować dla problemu sortowania. Niech $P = (X, \preceq)$ będzie *zbiorem częściowo uporządkowanym*, czyli zbiorem X wyposażonym w relację częściowego porządku $\preceq \subseteq X \times X$, spełniającą warunki:

(zwrotność) $x \preceq x$ dla każdego $x \in X$;
(przechodniość) jeśli $x \preceq y$ i $y \preceq z$, to $x \preceq z$ dla każdych $x, y, z \in X$;
(antysymetryczność) jeśli $x \preceq y$ i $y \preceq x$, to $x = y$ dla każdych $x, y \in X$.

Zbiorem *liniowo uporządkowanym* nazywamy zbiór częściowo uporządkowany, w którym każde dwa elementy są porównywalne, czyli spełniony jest dodatkowy warunek:

(spójność) $x \preceq y$ lub $y \preceq x$ dla każdych $x, y \in X$.

Rozszerzeniem liniowym zbioru częściowo uporządkowanego $P = (X, \preceq_P)$ nazywamy każdy zbiór liniowo uporządkowany $Q = (X, \preceq_Q)$, który zachowuje relację porządku, czyli $\preceq_P \subset \preceq_Q$. Rozszerzenie liniowe utożsamiamy z permutacją elementów zbioru X.

W dalszym ciągu będziemy rozważać tylko zbiory skończone. Dla przykładu niech $X = \{a, b, c\}$. Totalnym nieporządkiem na tym zbiorze jest zbiór częściowo uporządkowany $P_0 = (X, \preceq_0)$, taki że nie zachodzi $x \preceq_0 y$ dla dowolnych różnych $x, y \in X$. P_0 ma sześć rozszerzeń liniowych:

$$(a, b, c), (a, c, b), (b, a, c), (b, c, a), (c, a, b), (c, b, a).$$

Sortowanie zbioru częściowo uporządkowanego polega na zadawaniu pytań o relację między jego elementami w celu wybrania jednego z jego rozszerzeń liniowych. Niech $P_1 = (X, \preceq_1)$ będzie zbiorem częściowo uporządkowanym, który chcemy posortować. Jeśli na pytanie o elementy x i y uzyskamy odpowiedź, że $x \preceq y$, to rozszerzamy relację $\preceq_1$ o tę informację. Wynikiem jest nowa relacja $\preceq_2$ i nowy zbiór częściowo uporządkowany $P_2 = (X, \preceq_2)$. Formalnie relacja $\preceq_2$ jest *domknięciem przechodnim* relacji $\preceq_1$ uzupełnionej o porównanie $x \preceq y$.

Domknięcie przechodnie relacji dwuargumentowej ϱ na zbiorze X jest to najmniejsza (w sensie inkluzji) relacja przechodnia ϱ' na zbiorze X, taka że $\varrho \subseteq \varrho'$.

Sortowanie kończy się, gdy w wyniku zadawania kolejnych pytań otrzymamy zbiór liniowo uporządkowany. Oczywiście, nie ma sensu zadawanie pytania, na które znamy odpowiedź, czyli pytania o relację między elementami x i y, jeśli $x \preceq_1 y$ lub $y \preceq_1 x$. Dla zachowania ścisłości przyjmujemy, że po zadaniu takiego pytania relacja nie zmienia się.

Wróćmy do naszego przykładu trójelementowego zbioru X i totalnego nieporządku na nim. Chcąc go posortować, możemy zadać pytanie o relację między elementami a i b. Jeśli odpowiedzią jest $a \preceq b$, to nadal możliwe są rozszerzenia liniowe

$$(a, b, c), (a, c, b), (c, a, b).$$

Jeśli natomiast odpowiedź brzmi $b \preceq a$, to pozostają rozszerzenia liniowe

$$(b, a, c), (b, c, a), (c, b, a).$$

Zauważmy, że wykonanie porównania dzieli zbiór rozszerzeń liniowych na dwa rozłączne podzbiory. Niech $e(P)$ oznacza liczbę rozszerzeń liniowych zbioru częściowo uporządkowanego. Łatwo zauważyć, że dla dowolnego zbioru częściowo uporządkowanego P jeśli P' i P'' oznaczają rozszerzenie P odpowiednio o warunek $x \preceq y$ i $y \preceq x$, to $e(P) = e(P') + e(P'')$ i przynajmniej jedna z wartości $e(P')$ lub $e(P'')$ jest nie mniejsza niż $\frac{1}{2}e(P)$.

Zastanówmy się, jaka jest minimalna liczba porównań $C(P)$ potrzebna i zawsze wystarczająca do posortowania danego zbioru częściowo uporządkowanego P. Gdyby było tak, jak w powyższym przykładzie, że zawsze możemy podzielić zbiór rozszerzeń liniowych na równoliczne podzbiory, to wystarczyłoby $\lceil \log_2 e(P) \rceil$ porównań. Niestety, nie zawsze jest to możliwe. Rozważany przykład pokazuje, że możemy uzyskać trójelementowy zbiór rozszerzeń liniowych i wtedy najlepszy możliwy do uzyskania podział jest w stosunku $1:2$. Okazuje się, że najprawdopodobniej jest to przypadek najbardziej pesymistyczny. Mówi o tym hipoteza 1/3-2/3 sformułowana wiele lat temu niezależnie przez Kislicyna, Fredmana i Liniala.

Hipoteza 1. *Dla dowolnego skończonego zbioru częściowo uporządkowanego P, który nie jest liniowo uporządkowany, zawsze możemy wskazać dwa elementy, takie że w wyniku ich porównania (niezależnie od wyniku tego porównania) otrzymujemy rozszerzenie R, dla którego zachodzi*

$$\frac{1}{3}e(P) \leqslant e(R) \leqslant \frac{2}{3}e(P).$$

Powyższą hipotezę udowodniono dla wielu przypadków szczególnych, co pozwala wierzyć w jej prawdziwość. Jednak w ogólnym przypadku nadal pozostaje jednym z ważniejszych problemów otwartych teorii zbiorów częściowo uporządkowanych. Prawdziwość tej hipotezy implikuje możliwość posortowania zbioru częściowo uporządkowanego P za pomocą maksymalnie $\lceil \log_{1,5} e(P) \rceil$ porównań (przypomnijmy, że $\log_{1,5} n > \log_2 n$ dla $n > 1$).

Czy można lepiej? Wydaje się, że tak. Autor tego artykułu sformułował kilka lat temu hipotezę złotego podziału dla zbiorów częściowo uporządkowanych.

[*]Instytut Informatyki, Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

Hipoteza 2. *Dla dowolnego skończonego zbioru częściowo uporządkowanego P, który nie jest liniowo uporządkowany, zawsze możemy wskazać dwa kolejne porównania, takie że niezależnie od ich wyniku otrzymujemy kolejno rozszerzenia R i S, dla których zachodzi*

$$e(P) \geqslant e(R) + e(S).$$

Hipoteza złotego podziału jest bardziej ogólna od hipotezy 1/3-2/3. Innymi słowy, prawdziwość hipotezy złotego podziału implikuje prawdziwość hipotezy 1/3-2/3, gdyż każdy kontrprzykład dla hipotezy 1/3-2/3 jest również kontrprzykładem dla hipotezy złotego podziału. Załóżmy, że P jest kontrprzykładem dla hipotezy 1/3-2/3. Wtedy dla każdego porównania na P jeden z jego wyników daje nierówność $e(R) > \frac{2}{3}e(P)$. Oczywiście, dla każdego porównania na R możemy wskazać jego wynik, taki że $e(S) \geqslant \frac{1}{2}e(R)$. Stąd $e(R) + e(S) \geqslant \frac{3}{2}e(R) > e(P)$ i otrzymaliśmy kontrprzykład dla hipotezy złotego podziału.

Jak można się domyślić, hipotezy złotego podziału również nie udało się udowodnić w ogólności, ale dowiedziono jej w tak wielu przypadkach szczególnych, że mamy podstawy, aby wierzyć w jej prawdziwość.

Zmierzając do finału, zobaczmy dwa proste twierdzenia, które uzasadniają nazwę hipotezy. Niech F_n będzie n-tą liczbą Fibonacciego, rozpoczynając od $F_1 = F_2 = 1$, a $\varphi = (1+\sqrt{5})/2$ stałą złotej proporcji.

Twierdzenie 1. *Jeśli hipoteza o złotym podziale jest prawdziwa i zachodzi $e(P) < F_{n+3}$, to $C(P) \leqslant n$.*

Dowód prowadzimy przez indukcję. Dla $n = 0$ z założenia wynika, że $e(P) < F_3 = 2$, czyli zbiór P ma tylko jedno rozszerzenie liniowe, więc jest już posortowany. Dla $n = 1$ zachodzi $e(P) < F_4 = 3$, czyli P ma co najwyżej dwa rozszerzenia liniowe, które można rozróżnić jednym porównaniem.

Załóżmy teraz, że twierdzenie jest prawdziwe dla $n-1$ i $n-2$. Niech $e(P) < F_{n+3}$. Wybieramy dwa porównania, aby spełnić nierówność $e(P) \geqslant e(R) + e(S)$, gdzie R i S oznaczają zbiory częściowo uporządkowane otrzymane odpowiednio po pierwszym i drugim porównaniu. Z własności liczb Fibonacciego wynika, że zachodzi przynajmniej jedna z nierówności $e(R) < F_{n+2}$ lub $e(S) < F_{n+1}$. Z założenia indukcyjnego w pierwszym przypadku możemy posortować R za pomocą co najwyżej $n-1$ porównań, a w drugim posortować S za pomocą co najwyżej $n-2$ porównań. Zatem w obu przypadkach możemy dokończyć sortowanie P tak, aby nie przekroczyć n porównań.

Twierdzenie 2. *Jeśli hipoteza o złotym podziale jest prawdziwa, to*

$$\sup_P \frac{C(P)}{\log_\varphi e(P)} = 1.$$

W powyższym twierdzeniu kres górny jest po wszystkich skończonych zbiorach częściowo uporządkowanych, których dotyczy hipoteza złotego podziału, czyli skończonych i nieuporządkowanych liniowo. Niech P będzie takim zbiorem częściowo uporządkowanym. Niech n będzie liczbą porównań wymaganych do jego posortowania ($n \geqslant 1$ i nie można go posortować za pomocą $n-1$ porównań). Z poprzedniego twierdzenia wynika, że $e(P) \geqslant F_{n+2}$. Wiadomo, że ciąg Fibonacciego spełnia nierówność $F_{n+2} > \varphi^n$ dla $n \geqslant 1$, zatem

$$\frac{C(P)}{\log_\varphi e(P)} \leqslant \frac{n}{\log_\varphi F_{n+2}} < 1.$$

Linial wskazał ciąg zbiorów częściowo uporządkowanych L_n, nazywanych drabinami, takich że L_n ma n elementów, $C(L_n) = n-1$ oraz $e(L_n) = F_{n+1}$. Rysunek u dołu strony pokazuje diagramy Hassego kilku początkowych zbiorów L_n i pozwala odgadnąć regułę ich konstrukcji oraz genezę ich nazwy.

Diagram Hassego jest to sposób przedstawienia zbioru częściowo uporządkowanego $(X, \preceq)$ za pomocą grafu nieskierowanego. Elementy przedstawia się jako węzły, a relacje między elementami zaznacza się, łącząc te węzły krawędziami. Przy czym rysuje się tylko niezbędne krawędzie. Jeśli $x \preceq y$, to element x rysuje się niżej niż element y, a krawędź między nimi rysuje się tylko wtedy, gdy nie istnieje taki element $z \in X \setminus \{x, y\}$, że $x \preceq z \preceq y$.

Ponieważ

$$F_n = \frac{\varphi^n - (1-\varphi)^n}{\sqrt{5}},$$

więc mamy

$$\lim_{n\to\infty} \frac{C(L_n)}{\log_\varphi e(L_n)} = 1.$$

Z powyższych rozważań wnioskujemy, że postulowanego ograniczenia tego $C(P) \leqslant \log_\varphi e(P)$ nie da się już poprawić. Równość zachodzi wtedy i tylko wtedy, gdy P jest liniowo uporządkowany – wtedy $e(P) = 1$ i $C(P) = 0$. Zauważmy też, że $\log_{1,5} n > \log_\varphi n > \log_2 n$ dla $n > 1$. Nieco nieformalnie można powiedzieć, że podczas sortowania każde porównanie może zmniejszyć średnio liczbę rozszerzeń liniowych przynajmniej o współczynnik złotej proporcji. Tak właśnie jest dla drabin. Chcąc posortować drabinę L_n, należy porównać dwa jej elementy maksymalne. W wyniku zbiór jej rozszerzeń liniowych o liczności F_{n+1} zostaje podzielony na dwa podzbiory o liczności odpowiednio F_n i F_{n-1}, a problem redukuje się do posortowania odpowiednio drabiny L_{n-1} lub L_{n-2}.

Element maksymalny w zbiorze częściowo uporządkowanym $(X, \preceq)$ jest to taki element $x \in X$, że nie istnieje element $y \in X \setminus \{x\}$, dla którego $x \preceq y$.