

# O przydziale miejsc na przedmiotach

Oskar SKIBSKI\*

\*Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

Podobno zdarzało się nawet, że studenci z wysoką średnią rejestrowali się na jakiś przedmiot „dla kogoś”, a potem na papierowych druczkach w Sekcji Studenckiej oddawali mu to miejsce.

W dyskusjach i opracowaniu tej metody uczestniczyło sporo osób: Marcin Engel, Agata Janowska, Piotr Kępczyński, Paulina Kubera, Adam Malinowski, Konrad Mocarski, Jakub Piotrowicz, Barbara Próchniak, Oskar Skibski i Piotr Skowron (kolejność alfabetyczna).

Kilka miesięcy temu Samorząd Studentów Wydziału Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego zaproponował, aby zmienić sposób, w jaki odbywa się rejestracja studentów na przedmioty. Do tej pory, w uproszczeniu, działała ona tak: każdy student wybierał przedmioty, na które chciałby się zapisać; następnie po kolei rozpatrywani byli studenci, w kolejności malejącej średniej ocen, i każdemu przydzielane były miejsca na wszystkich przedmiotach, które wybrali i na które były jeszcze wolne miejsca. Studenci nie byli zbyt zadowoleni z tego systemu, i trudno im się dziwić. Jednym z problemów było to, że student nie wiedział, ile przedmiotów powinien wybrać, bo nie miał pewności, ile z nich dostanie, a ileś zaliczyć musi. Większym problemem było jednak to, że system ten bardzo mocno premiował osoby z wysoką średnią. Można argumentować, że student z najwyższą średnią powinien mieć pierwszeństwo rejestracji na pewną liczbę przedmiotów przed studentem ze średnią najniższą. Ale czy powinien też mieć pierwszeństwo na wszystkich przedmiotach przed studentem ze średnią tylko odrobinę niższą?

Pierwszym naturalnym krokiem do uzdrowienia systemu było wprowadzenie limitu na liczbę przedmiotów, jakie może dostać jeden student, oraz zmiana tego, co deklarują studenci – lepiej, aby podawali oni kolejność przedmiotów, na które chcą się zarejestrować, od najbardziej pożądanego do najmniej. Trudniejszym problemem jest jednak wyrównanie szans studentów. Najrówniej byłoby oczywiście, gdyby w kółko każdemu studentowi przydzielał jeden przedmiot. Ten system, znany na przykład z wyboru drużyn na lekcjach wf, nazywa się *Round Robin* (lub rzadziej, ale po polsku, *algorytmem karuzelowym*). Jednak czy tak powinna działać rejestracja na przedmioty? Czy ciężko wypracowana średnia nie powinna odgrywać większej roli? Czy dałoby się zrobić tak, aby osoby ze zbliżoną średnią były traktowane tak równo, jak się da, ale większa różnica w średnich przekładała się na lepszą sytuację?

Pewnie, że się da! W wyniku wielu rozmów i negocjacji powstał algorytm, który ze względu na swoją uderzającą sprawiedliwość nazwany został *Robin Hoodem*.

## Przydział miejsc na przedmiotach (*Robin Hood*)

**Krok 1:** Każdy student podaje ranking przedmiotów, na jakie chciałby się zapisać.

**Krok 2:** Umieszczamy wszystkich studentów na wspólnej liście i sortujemy malejąco względem średniej ocen.

**Krok 3:** Bierzemy pierwszą osobę z listy i przydzielamy jej przedmiot, na którym zależy jej najbardziej spośród tych, na których są jeszcze wolne miejsca (jeżeli jeszcze nie osiągnęła pewnego z góry ustalonego limitu).

**Krok 4:** Zmniejszamy średnią wybranej osobie o stałą  $c = 0,5$  i przesuwamy ją odpowiednio w dół na liście.

**Krok 5:** Powtarzamy kroki 3 i 4, aż skończą się nam miejsca na przedmiotach lub nikt nie będzie chciał już żadnego z dostępnych przedmiotów.

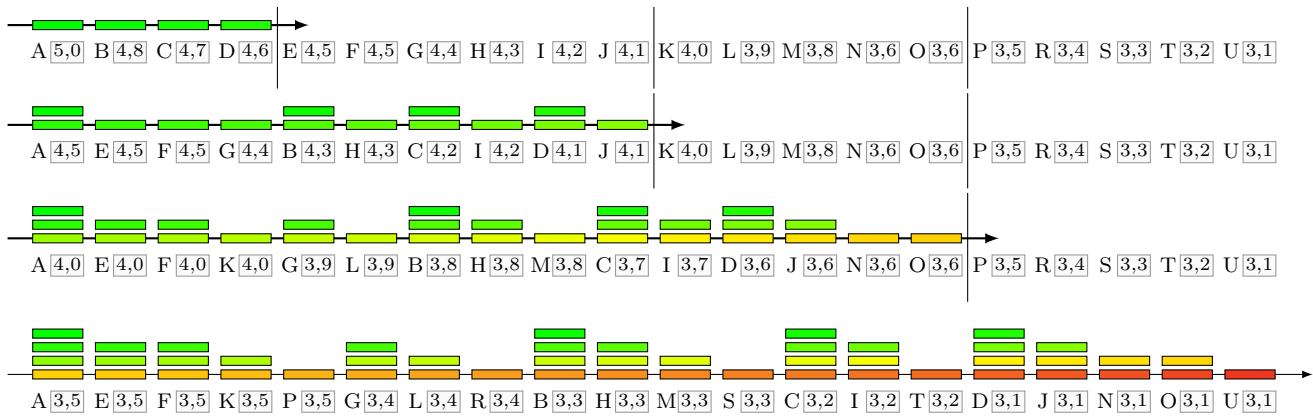
HURRA  
JESTEM!



Ile powinna wynosić stała  $c$ , jest kwestią do dyskusji. Łatwo zauważyć, że gdyby zmienić ją na 0, to uzyskalibyśmy obecny system, a dla 5 to zwykły Round Robin.

A co ta stała oznacza? Ustalmy dwóch studentów i popatrzmy, jak przeplatają się ich wybory. Jeżeli dwie osoby mają różnicę średnich mniejszą niż  $c$ , to są traktowane tak równo, jak się da – będą wybierać na zmianę, zaczynając od tej z wyższą średnią. Jeżeli z kolei ta różnica jest pomiędzy  $k \cdot c$  a  $(k + 1) \cdot c$ , to student z wyższą średnią najpierw wybierze  $k$  przedmiotów, a potem będą wybierać na zmianę, zaczynając od tego z wyższą średnią. Na przykład dla najwyższej średniej 5,0 i najniższej 3,1, najlepszy student wybierze 4 przedmioty przed najgorszym.

Możemy się spierać, czy metoda jest fajna, czy nie. W szanowanym czasopiśmie, jakim jest *Delta*, skupimy się jednak na konkretnych, czyli zbadamy, jakie ma teoretyczne własności. Ale najpierw ilustracja.



Ilustracja prezentuje kolejność wyborów w Robin Hoodzie, a kolejne wiersze odpowiadają przydziałowi miejsc na przedmiotach po 4, 14, 29 i 49 krokach procedury. Kolorowe prostokąty symbolizują przydzielane miejsca, a kolor odpowiada kolejności, w której miejsce zostało przyznane – pierwsze miejsca są zielone, a ostatnie czerwone. Jak widzimy, studenci z najwyższą średnią ( $> 4,5$ ) wybierają po pierwszym przedmiocie, po czym zmniejszamy im średnią o 0,5. W ten sposób „wtasowują się” oni w studentów ze średnią pomiędzy 4,0 a 4,5, i razem z nimi wybierają po kolejnym przedmiocie.

Aby opowiedzieć o własnościach Robin Hooda, wprowadźmy trochę oznaczeń. Załóżmy, że mamy  $n$  ponumerowanych studentów. Przyjmijmy, że każdy student ma listę rankingową wszystkich przedmiotów, bez remisów. Dla dwóch przedmiotów  $a, b$  będziemy pisać  $a \succ_i b$ , jeżeli student  $i$  woli przedmiot  $a$  od  $b$ . Dla uproszczenia zignorujemy limit przedmiotów, jaki może mieć jeden student: nie zmienia on wiele, a komplikuje analizę. Będziemy więc każdemu studentowi przydzielać jak najwięcej miejsc.

Między studentów będziemy rozdzielać pewien zbiór miejsc na przedmiotach. Dla ułatwienia ten zbiór będzie po prostu multizbiorem przedmiotów, czyli „zbiorem”, w którym przedmiot pojawia się tyle razy, ile jest na nim miejsc. Szukamy pewnego *podziału* tych miejsc, czyli listy  $(A_i) = (A_1, \dots, A_n)$  takiej, że  $A_1, \dots, A_n$  są zbiorami przedmiotów dla studentów  $1, \dots, n$  (nie multizbiorami!), które połączone dadzą multizbiór wszystkich miejsc.

Co możemy powiedzieć o preferencjach studenta dla zbiorów przedmiotów? Aby uniknąć spekulacji, czy lepiej mieć dwa średnie przedmioty, czy jeden fajny i jeden słaby, będziemy porównywać tylko niektóre zbiory. Powiemy, że osobie  $i$  podoba się zbiór  $A$  co najmniej tak samo jak  $B$  (oznaczane  $A \succeq_i B$ ), jeżeli  $|A| \geq |B|$  oraz najbardziej lubiany przedmiot z  $A$  jest co najmniej tak lubiany jak ulubiony z  $B$ , drugi z  $A$  co najmniej tak lubiany jak drugi z  $B$  itd. Łatwo zauważyć, że  $A \succeq_i B$  oraz  $B \succeq_i A$  implikuje  $A = B$ . Będziemy więc pisać  $A \succ_i B$ , jeżeli  $A \succeq_i B$  i  $A \neq B$ .

Jedną z podstawowych własności dobrego podziału jest to, że nie da się go łatwo poprawić, to znaczy tak pozmienić, aby nikt na tym nie stracił, a ktoś zyskał.

Podział  $(A_i)$  jest *optymalny w sensie Pareto*, jeżeli nie istnieje podział  $(B_i)$  taki, że  $B_i \succeq_i A_i$  dla każdego studenta  $i$  oraz  $B_i \succ_i A_i$  dla przynajmniej jednego studenta.

Czy podział uzyskany przez Robin Hooda jest w tym sensie optymalny? Tak. Można powiedzieć nawet więcej – dowolny podział uzyskany przez kolejne przydzielanie studentom miejsc na ich ulubionych dostępnych przedmiotach jest optymalny, niezależnie od kolejności, w jakiej będzie się te osoby wybierać. Aby to zobaczyć, wystarczy po kolei rozpatrywać przyznane miejsca. Czy student wybierający jako pierwszy może preferować podział, w którym nie ma przedmiotu, który dostał w pierwszej kolejce? Oczywiście nie. A spośród takich podziałów, czy student, który wybiera jako drugi, może preferować podział bez przedmiotu, który dostał w drugiej kolejce? Nie – jeżeli nie jest to jego ulubiony

Przykładowo dla dwóch przedmiotów  $m, p$  (matematyka i polski) możemy mieć multizbiór  $O = \{m, m, m, p, p\}$  oznaczający 3 miejsca na matematyce i 2 na polskim.

Formalnie,  $A \succeq_i B$ , jeżeli  $|A| \geq |B|$  i możemy tak uporządkować przedmioty  $A = \{a_1, \dots, a_h\}$  i  $B = \{b_1, \dots, b_\ell\}$ , aby zachodziło  $a_j = b_j$  lub  $a_j \succ_i b_j$  dla każdego  $j \in \{1, \dots, \ell\}$ .

przedmiot, to znaczy, że na ulubionym było tylko jedno miejsce, którego nie może dostać, nie krzywdząc w ten sposób pierwszego studenta. I tak dalej.

**Drunk Robin:** Wybieramy dowolnego studenta i przydzielamy mu jego ulubiony dostępny przedmiot. Powtarzamy, dopóki są jeszcze jakieś miejsca.

Nazwijmy taki algorytm przydziału Drunk Robinem (definicja na marginesie). Okazuje się, że wszystkie optymalne podziały da się nim uzyskać!

**Twierdzenie.** *Podział jest optymalny w sensie Pareto wtedy i tylko wtedy, gdy da się go uzyskać algorytmem Drunk Robin.*

Jeżeli w grafie wszystkie wierzchołki mają krawędzie wychodzące, to zawsze znajdziemy w nim cykl, podążając uparcie za strzałkami wychodzącymi – studentów (na świecie) mamy skończoną liczbę, więc w końcu ktoś się powtórzy.

*Dowód.* Uzasadniliśmy już, że Drunk Robin daje tylko podziały optymalne. Załóżmy, że mamy podział optymalny i rozpatrzmy graf, w którym wierzchołkami są studenci, a skierowana krawędź  $(i, j)$  istnieje, jeżeli  $j$  posiada jakiś przedmiot, który  $i$  woli od wszystkich swoich przedmiotów. Gdyby ten graf miał cykl, to podział nie byłby optymalny – każdy student mógłby dostać bardziej lubiany przedmiot w zamian za któryś mniej lubiany. Skoro graf nie ma cyklu, to musi istnieć wierzchołek bez krawędzi wychodzących, czyli który nikomu nie zazdrości żadnego przedmiotu. Możemy mu zatem pozwolić wybierać jako pierwszemu. Następnie kasując jego ulubiony przedmiot z jego zbioru oraz jedno miejsce ze zbioru dostępnych miejsc, dostaniemy mniejszy podział, który też jest optymalny: gdyby nie był i istniałby lepszy podział, po dodaniu skasowanego przedmiotu dostalibyśmy podział lepszy niż oryginalny, przecząc jego optymalności. Postępując dla niego i kolejnych mniejszych podziałów analogicznie, otrzymamy kolejność, w której każdy student wybiera ulubiony dostępny przedmiot.  $\square$

Podziały uzyskane z Drunk Robina mogą być jednak bardzo niesprawiedliwe. W jaki sposób możemy opisać fakt, że podział jest sprawiedliwy? Nie możemy wymagać, aby wszyscy dostali miejsce na ulubionym przedmiocie. Moglibyśmy jednak chcieć, aby każda osoba miała zbiór nie więcej niż o jeden przedmiot gorszy niż każda inna. O tym właśnie mówi następująca właściwość:

Podział  $(A_i)$  jest *wolny od zazdrości z dokładnością do jednego przedmiotu (EF1)*, jeżeli dla dowolnych dwóch studentów  $i, j$  zachodzi  $A_j = \emptyset$  albo istnieje przedmiot  $p \in A_j$ , taki że  $A_i \succeq_i A_j \setminus \{p\}$ .

**Round Robin:** Układamy studentów w listę. Bierzemy pierwszego, przydzielamy mu jego ulubiony dostępny przedmiot i przesuwamy go na koniec listy. Powtarzamy, dopóki są jeszcze jakieś miejsca.

Rozważmy 2 studentów, zbiór miejsc  $\{a, b, c, d\}$  (na każdym przedmiocie jedno miejsce), podział  $A_1 = \{a, b\}$ ,  $A_2 = \{c, d\}$  i preferencje:  $a \succ_1 b \succ_1 c \succ_1 d$  oraz  $b \succ_2 c \succ_2 d \succ_2 a$ . Podział ten jest optymalny w sensie Pareto i spełnia EF1, ale nie da się go uzyskać algorytmem Round Robin.

Dla modelu, w którym nie ma limitu przedmiotów, jakie może mieć jeden student, definicja Robin Hooda wygląda tak:

**Robin Hood:** Wybieramy studenta z najwyższą średnią, przydzielamy mu jego ulubiony dostępny przedmiot i zmniejszamy mu średnią o  $c = 0,5$ . Powtarzamy, dopóki są jeszcze jakieś miejsca.

Łatwo zobaczyć, że podziały uzyskane Round Robinem spełniają tę własność: pierwszy przedmiot wybrany przez studenta  $i$  jest dla niego lepszy lub taki sam jak drugi wybrany przez studenta  $j$ , drugi wybrany przez  $i$  jest lepszy lub taki sam jak trzeci wybrany przez  $j$  itd. Co ciekawe, istnieją jednak podziały optymalne w sensie Pareto, które spełniają EF1, a nie da się ich uzyskać algorytmem Round Robin. Przykład znajduje się na marginesie.

A czy podziały uzyskane Robin Hoodem spełniają EF1? Nie, ale to dobrze – chcemy przecież, aby wyższa średnia przełożyła się na znacząco lepszy zbiór przedmiotów. Niech  $p_i$  będzie pewnym priorytetem, czyli u nas średnią studenta  $i$ . Możemy zdefiniować następujący wariant własności EF1:

Podział  $(A_i)$  jest *wolny od zazdrości z dokładnością do jednego przedmiotu, uwzględniając priorytety (EF1-p)*, jeżeli dla dowolnych dwóch studentów  $i, j$  zachodzi:

- jeżeli  $(p_j - p_i)/c \in [\ell, \ell + 1)$  dla  $\ell \in \mathbb{N}$ , to istnieje  $P \subseteq A_j$ ,  $|P| \leq \ell + 1$  taki, że  $A_i \succeq_i A_j \setminus P$ ,
- jeżeli  $(p_j - p_i)/c \in (\ell, \ell + 1]$  dla  $\ell \in \mathbb{N}$ , to istnieje  $P \subseteq A_j$ ,  $|P| \leq \ell$  taki, że  $A_j \setminus P \succeq_j A_i$ .

Łatwo sprawdzić, że gdy priorytety są równe, własność ta sprowadza się do EF1. Przy priorytetach spełnialność wynika z kolejności, w jakiej studenci wybierają przedmioty: dla  $(p_j - p_i)/c \in [\ell, \ell + 1)$  student  $j$  wybiera  $\ell$  lub  $\ell + 1$  przedmiotów przed studentem  $i$ , więc po wyrzuceniu  $\ell + 1$  ulubionych przedmiotów studentowi  $j$  student  $i$  będzie wolał swój zbiór.

Nasz algorytm zastosowaliśmy do przydziału miejsc na przedmioty, ale można go też stosować w codziennym życiu, np. do podziału czekoladek z bombonierki. Starszym Czytelnikom sugerujemy przyjąć za priorytet wiek, tylko ile powinno wynosić  $c$ ? Nie będziemy osądzać, jeżeli zostaniecie przy  $c = 0,5$ .