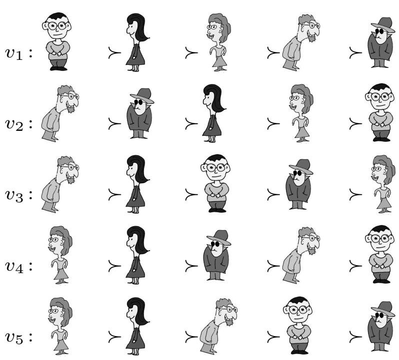
Proporcjonalność bez partii politycznych

*Piotr SKOWRON**

* Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

Rozważmy następujący scenariusz: n wyborców pragnie wybrać komitet, czyli podzbiór kandydatów o ustalonej liczebności. Oznaczmy zbiór wyborców jako $N = \{1, 2, \ldots, n\}$, zbiór kandydatów jako $C = \{c_1, \ldots, c_m\}$ (w szczególności n i m to, odpowiednio, liczba wyborców i liczba kandydatów), a rozmiar komitetu, który chcemy wyłonić jako k (przy czym $k \leqslant m$). Na przykład wyobraźmy sobie, że wyborcy chcą wybrać parlament o rozmiarze k w małym kraju, w którym nie zaistniała jeszcze koncepcja partii politycznych (w związku z czym muszą głosować bezpośrednio na kandydatów, a nie na partie polityczne), albo że pracownicy pewnej firmy lub organizacji chcą wybrać k osób, które będą ich reprezentować w związkach zawodowych. Załóżmy, że każdy z wyborców głosuje, wskazując podzbiór kandydatów, których uważa za akceptowalnych: dla wyborcy $i \in N$ taki podzbiór akceptowalnych kandydatów będziemy oznaczać przez $A(i)$. Pokażemy, że wskazanie „sprawiedliwego" sposobu przeprowadzenia takich wyborów jest nieoczywiste.

Innym klasycznym modelem rozważanym w ramach teorii wyboru społecznego jest model oparty o rankingi. Wtedy wymagamy, aby wyborcy uszeregowali kandydatów od najbardziej do najmniej preferowanego. W takim modelu przykład preferencji wyborców może wyglądać następująco:

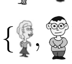


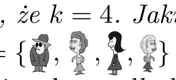
Model, w którym wyborcy wskazują kandydatów, których akceptują, jest jednak szczególnie atrakcyjny. W takim systemie głosowania wyborcom stosunkowo łatwo jest oddać głos, podczas gdy uszeregowanie, nawet tylko kilkunastu kandydatów, jest dużo trudniejsze.

Przykład 1. *Przyjmijmy, że $n = 800$, $m = 9$, a głosy wyborców są następujące:*

400 wyborców akceptuje: {}

200 wyborców akceptuje: {}

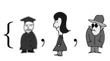
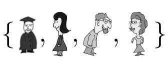
200 wyborców akceptuje: {}

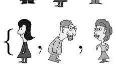
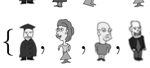
Przypuśćmy, że $k = 4$. Jaki komitet powinien zostać wybrany w tym przypadku? Komitet $S =$ {} *wydaje się być rozsądnym wyborem*[1]*. Taki komitet jest* proporcjonalny *– dla każdej z trzech grup wyborców, liczba wybranych kandydatów, których ci wyborcy aprobują, jest proporcjonalna do liczby wyborców w tej grupie.*

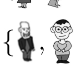
[1] Oczywiście nie jest to jedyny proporcjonalny komitet. Każdy komitet, który będzie składał się z dwóch kandydatów akceptowanych przez pierwszą grupę 400 wyborców, jednego kandydata akceptowanego przez drugą grupę 200 wyborców i jednego kandydata akceptowanego przez ostatnią grupę wyborców, będzie równie dobry. Na potrzeby naszej analizy załóżmy, że mamy pewną regułę rozstrzygania remisów, używaną, gdy istnieje kilka optymalnych komitetów.

Preferencje wyborców w przykładzie 1 mają pewną specyficzną strukturę: dla każdych dwóch wyborców ich zbiory akceptowalnych kandydatów są takie same lub rozłączne. Pozwala to pogrupować wyborców i kandydatów w taki sposób, że wskazanie proporcjonalnego komitetu jest proste i intuicyjne. Co jednak, gdy mamy do czynienia z bardziej złożonymi i „chaotycznymi" preferencjami?

Przykład 2. *Rozważmy następujące preferencje wyborców:*

v_1 : {} v_2 : {}

v_3 : {} v_4 : {}

v_5 : {} v_6 : {}

v_7 : {} v_8 : {}.

Przypuśćmy, że chcemy wybrać proporcjonalny komitet rozmiaru $k = 4$. W tym wypadku odpowiedź nie jest już tak oczywista.

Wyżej opisany model był badany – między innymi – pod koniec dziewiętnastego wieku przez duńskiego astronoma i matematyka Thorvalda Nicolai Thielego. Thiele zaprojektował regułę obecnie nazywaną PAV (*Proportional Approval Voting*). Dla $r \in \mathbb{N}$, niech $\mathrm{H}(r)$ oznacza r-tą liczbę harmoniczną, tzn. $\mathrm{H}(r) = \sum_{i=1}^{r} 1/i$. Reguła PAV wybiera komitet S, który maksymalizuje wartość $\mathrm{pkt}(S) = \sum_{i \in N} \mathrm{H}(|A(i) \cap S|)$. Innymi słowy, dla każdego wyborcy $i \in N$ i dla każdego komitetu $S \subseteq C$ ($|S| = k$) analizujemy, ilu kandydatów z komitetu S wyborca i akceptuje ($|A(i) \cap S|$); bierzemy liczbę harmoniczną o tym indeksie z tej wartości ($\mathrm{H}(|A(i) \cap S|)$) i traktujemy ją jako wynik punktowy, który wyborca i przypisuje komitetowi S. Dla każdego komitetu liczymy sumę punktów przypisanych temu komitetowi przez wszystkich wyborców i wybieramy ten komitet S, którego wynik punktowy $\mathrm{pkt}(S)$ jest najwyższy. Na pierwszy

Thorvald Nicolai Thiele (1838–1910), profesor astronomii i dyrektor obserwatorium astronomicznego na Uniwersytecie w Kopenhadze. Jego najbardziej znane prace dotyczą matematyki, w szczególności statystyki.

Koniec dziewiętnastego wieku był szczególnie korzystnym okresem dla naukowców zajmujących się projektowaniem systemów wyborczych. Na przykład Szwecja wprowadzała w tym czasie prawo wyborcze dla kobiet; w związku z tym partia konserwatywna spodziewała się utraty większości parlamentarnej i skłonna była rozważać zmianę dotychczasowej ordynacji wyborczej. Inny, równie ciekawy, system wyborczy został opracowany mniej więcej w tym samym czasie przez szwedzkiego matematyka Larsa Edvarda Phragména (1863–1937).

rzut oka nie jest jasne, dlaczego użyliśmy liczb harmonicznych w definicji reguły PAV[2]. Jak się jednak okaże, w tym przypadku jest to jak najbardziej uzasadnione.

W przykładzie 2 zastosowanie reguły PAV doprowadzi do wyboru komitetu $S = \{\ , \ , \ , \ \}$. Wyborca v_1 akceptuje jednego członka tego komitetu, więc przypisze mu $\mathrm{H}(1) = 1$ punkt; wyborca v_2 akceptuje trzech kandydatów z S, więc przypisze temu komitetowi $\mathrm{H}(3) = 1 + 1/2 + 1/3 = 11/6$ punktów. Sumarycznie komitet ten otrzyma następującą liczbę punktów:

$$\mathrm{pkt}(\{\ , \ , \ , \ \}) = 1 + 11/6 + 3/2 + 3/2 + 1 + 1 + 1 + 3/2 = 31/3.$$

PAV wybierze ten właśnie komitet, ponieważ wynik punktowy dla każdego innego 4-elementowego podzbioru kandydatów będzie mniejszy bądź równy 31/3.

Dlaczego regułę wyborczą PAV możemy uważać za proporcjonalną? Tu z pomocą przychodzi nam podejście aksjomatyczne: najpierw spróbujemy formalnie zdefiniować pewną matematyczną własność, która będzie mówiła, czy dana reguła wyborcza jest proporcjonalna, czy też nie; następnie sprawdzimy, czy PAV rzeczywiście spełnia tę własność.

Definicja (silna proporcjonalność). *Powiemy, że reguła wyborcza $\mathcal{R}$ spełnia* silną proporcjonalność*, jeśli dla każdego profilu preferencji, dla każdej liczby $\ell \in \mathbb{N}$ oraz dla każdej grupy wyborców $V \subseteq N$ takiej, że $|V| \geqslant \ell \cdot n/k$ oraz $|\bigcap_{i \in V} A(i)| \geqslant \ell$, reguła zwróci komitet S taki, że $\frac{1}{V} \cdot \sum_{i \in V} |A(i) \cap S| \geqslant \ell$.*

Zgodnie z powyższą definicją grupa $\ell \cdot n/k$ wyborców (czyli grupa, która stanowi ℓ/k część wszystkich wyborców) powinna mieć prawo decydować o ℓ członkach komitetu (czyli o ℓ/k części wybranego komitetu); np. grupa 30% wyborców powinna móc decydować o 30% członków komitetu. Zatem jeżeli grupa ma co najmniej $\ell \cdot n/k$ wyborców ($|V| \geqslant \ell \cdot n/k$) oraz jeżeli grupa ta zgadza się co do tego, że akceptuje pewnych ℓ kandydatów ($|\bigcap_{i \in V} A(i)| \geqslant \ell$), to średnio wyborcy z tej grupy powinni akceptować co najmniej ℓ członków wybranego komitetu ($\frac{1}{V} \cdot \sum_{i \in V} |A(i) \cap S| \geqslant \ell$).

Powyższa własność jest bardzo pożądana, ale nie istnieje reguła wyborcza, która by ją spełniała. Aby się o tym przekonać, spójrzmy na następujący przykład:

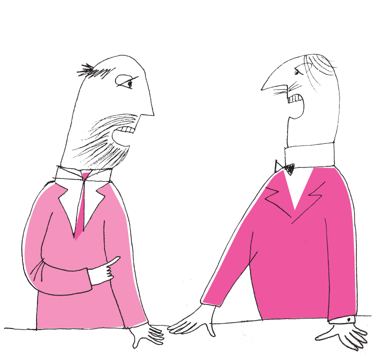
$v_1 : \{\ , \ \}$	$v_4 : \{\ , \ \}$	$v_7 : \{\ , \ \}$	$v_{10} : \{\ , \ \}$
$v_2 : \{\ \}$	$v_5 : \{\ \}$	$v_8 : \{\ \}$	$v_{11} : \{\ \}$
$v_3 : \{\ \}$	$v_6 : \{\ \}$	$v_9 : \{\ \}$	$v_{12} : \{\ \}$.

Przyjmijmy, że $k = 3$. W tym przykładzie grupa wyborców $v_1, \ldots, v_4$ stanowi 1/3 społeczeństwa i zgadza się co do kandydata . Wyborcy ci powinni zatem akceptować średnio co najmniej jednego członka komitetu, więc musi zostać wybrany jako członek zwycięskiego komitetu (gdybyśmy nie wybrali , to nawet przy wyborze i wyborcy $v_1, \ldots, v_4$ akceptowaliby średnio tylko 1/2 członków wybranego komitetu). Analogiczne rozumowanie zastosowane do grupy $v_4, \ldots, v_7$ pozwala nam stwierdzić, że również musi być członkiem zwycięskiego komitetu. I tu podobnie stwierdzamy, że oraz muszą zostać wybrani. Jest to niemożliwe, ponieważ możemy wybrać jedynie trzech kandydatów.

Możemy natomiast delikatnie osłabić powyższą własność: będziemy wymagać, aby grupa wyborców V, która spełnia warunki $|V| \geqslant \ell \cdot n/k$ i $|\bigcap_{i \in V} A(i)| \geqslant \ell$ (czyli grupa, której intuicyjnie powinno przysługiwać prawo wyboru ℓ członków komitetu), akceptowała średnio co najmniej $\ell - 1$ wybranych kandydatów. W definicji silnej proporcjonalności zmieniamy warunek $\frac{1}{V} \cdot \sum_{i \in V} |A(i) \cap S| \geqslant \ell$ na $\frac{1}{V} \cdot \sum_{i \in V} |A(i) \cap S| \geqslant \ell - 1$ i tę własność będziemy nazywać *słabą proporcjonalnością*.

[2]W regule PAV maksymalizujemy wartość zadowolenia wyborców. Wyborca będzie bardziej zadowolony, jeżeli wybory wygra chociaż jeden z akceptowanych przez niego kandydatów, niż gdyby miał nie wygrać żaden. Oczywiście będzie także bardziej zadowolony, jeżeli wygra dwóch kandydatów, których akceptuje, jeszcze bardziej, jeżeli wygra trzech takich kandydatów itd. Zadowolenie z każdym kolejnym akceptowanym kandydatem będzie rosło coraz wolniej. Jest wiele funkcji, poza funkcją harmoniczną, które spełniają tę własność. Wykażemy jednak, że funkcja harmoniczna jest w pewien sposób „specjalna” i użycie akurat tej funkcji powoduje, że system wyborczy jest proporcjonalny.

Aby reguła PAV mogła wyłonić zwycięski komitet, trzeba obliczyć wynik punktowy dla każdego możliwego komitetu. Takich komitetów jest $\binom{m}{k}$. Powoduje to pewne praktyczne ograniczenia: wydaje się, że ciężko użyć reguły PAV, zwłaszcza dla dużych wartości m i k. Współczesne prace z dziedziny informatyki pokazują, że potrafimy sobie dobrze radzić z tym problemem, stosując nowoczesne algorytmy heurystyczne i aproksymacyjne. Ponadto istnieją alternatywne systemy głosowania, które mają podobne własności co reguła PAV, a których obliczanie jest dużo prostsze, np. sekwencyjny wariant PAV lub sekwencyjna reguła Phragména.



Reguła PAV spełnia własność słabej proporcjonalności, co udowodnimy poniżej. Co więcej, gdybyśmy w regule PAV liczyli punkty komitetu, używając innej funkcji niż harmoniczna $\mathrm{H}(r) = \sum_{i=1}^{r} 1/i$, to taka reguła już nie spełniałaby własności słabej proporcjonalności[3]. Innymi słowy, to właśnie użycie liczb harmonicznych w definicji reguły PAV sprawia, że reguła ta jest proporcjonalna.

[3]Formalne wykazanie, że jakakolwiek reguła wyborcza licząca punkty przy użyciu innej funkcji niż funkcja harmoniczna $\mathrm{H}(r)$ nie spełnia słabej proporcjonalności, wymaga dłuższego dowodu. Zainteresowanego Czytelnika odsyłamy do pracy [1] (Twierdzenie 11). Pokażemy jednak, jak przeprowadzić takie rozumowanie w konkretnym przypadku. Załóżmy, że liczymy punkty, używając funkcji $g(r) = r/\mathrm{H}(r)$ zamiast $\mathrm{H}(r)$, czyli że używamy funkcji rosnącej szybciej niż funkcja harmoniczna. Rozważmy teraz wybory, w których 70 wyborców akceptuje kandydatów $c_1, \ldots, c_{10}$ i 30 wyborców akceptuje kandydatów $c_{11}, \ldots, c_{20}$. Niech $k = 10$. Łatwo sprawdzić, że komitet, który maksymalizuje wynik punktowy, składa się z 9 kandydatów spośród $c_1, \ldots, c_{10}$ i jednego kandydata spośród $c_{11}, \ldots, c_{20}$ – taki maksymalny wynik punktowy to $70 \cdot 9/\mathrm{H}(9) + 30 \cdot 1/\mathrm{H}(1) \approx 252{,}7$. Zatem w tym wypadku grupa 30 wyborców akceptuje średnio tylko jednego kandydata z wybranego komitetu. Własność słabej proporcjonalności wymaga, aby członkowie tej grupy akceptowali średnio co najmniej dwóch kandydatów z wybranego komitetu. Pokazuje to, że użycie funkcji $g(r) = r/\mathrm{H}(r)$ nie daje własności słabej proporcjonalności.

Twierdzenie ([2])**.** *Reguła PAV spełnia własność słabej proporcjonalności.*

Dowód. Przeprowadzimy dowód nie wprost. Załóżmy, że reguła PAV wybrała komitet S oraz że istnieje grupa wyborców V i liczba $\ell \geqslant 1$ taka, że:

(i) $|V| \geqslant \ell \cdot n/k$,
(ii) $|\bigcap_{i \in V} A(i)| \geqslant \ell$ oraz
(iii) $\frac{1}{V} \cdot \sum_{i \in V} |A(i) \cap S| < \ell - 1$.

Wykażemy, że z komitetu S możemy usunąć jednego kandydata, a w zamian dodać innego, tak że wynik punktowy komitetu $\mathrm{pkt}(S)$ wzrośnie. Będzie to sprzeczność z tym, że komitet S został wybrany przez regułę PAV; wszak PAV wybiera tylko te komitety, których wartość punktowa jest najwyższa.

Dla każdego wybranego kandydata $c \in S$, przez $\Delta(c)$ oznaczmy wartość, o którą zmniejszy się wynik punktowy S, gdy usuniemy z niego c. Formalnie $\Delta(c) = \mathrm{pkt}(S) - \mathrm{pkt}(S \setminus \{c\})$. Przez $N(c)$ oznaczmy zbiór wyborców, którzy akceptują kandydata c, $N(c) = \{i \in N : c \in A(i)\}$. Oszacujmy sumę $\sum_{c \in S} \Delta(c)$:

$$\begin{aligned}\sum_{c \in S} \Delta(c) &= \sum_{c \in S} (\mathrm{pkt}(S) - \mathrm{pkt}(S \setminus \{c\})) = \\ &= \sum_{c \in S} \sum_{i \in N} (\mathrm{H}(|A(i) \cap S|) - \mathrm{H}(|A(i) \cap (S \setminus \{c\})|)) = \\ &= \sum_{c \in S} \sum_{i \in N(c)} (\mathrm{H}(|A(i) \cap S|) - \mathrm{H}(|A(i) \cap S| - 1)) = \\ &= \sum_{c \in S} \sum_{i \in N(c)} \frac{1}{|A(i) \cap S|} = \sum_{i \in N} \sum_{c \in A(i) \cap S} \frac{1}{|A(i) \cap S|} = \\ &= \sum_{i \in N \colon |A(i) \cap S| \geqslant 1} \left(|A(i) \cap S| \cdot \frac{1}{|A(i) \cap S|} \right) \leqslant |N| = n.\end{aligned}$$

Zatem ponieważ $|S| = k$, istnieje kandydat $c_r \in S$, dla którego $\Delta(c_r) \leqslant n/k$.

Rozważmy teraz kandydata c_a, który jest akceptowany przez wszystkich wyborców z V oraz który nie został wybrany do zwycięskiego komitetu. Taki kandydat zawsze istnieje; jeżeli wszyscy kandydaci z $\bigcap_{i \in V} A(i)$ należeliby do S, to oznaczałoby, że wyborcy z V akceptują średnio co najmniej ℓ członków S; założyliśmy jednak, że $\frac{1}{V} \cdot \sum_{i \in V} |A(i) \cap S| < \ell - 1$. Niech $\Delta(c_a)$ oznacza wartość, o jaką zwiększy się wynik punktowy komitetu S, gdy dodamy do niego c_a, $\Delta(c_a) = \mathrm{pkt}(S \cup \{a\}) - \mathrm{pkt}(S)$. Mamy:

$$\begin{aligned}\Delta(c_a) &= \sum_{i \in N} (\mathrm{H}(|A(i) \cap (S \cup c_a)|) - \mathrm{H}(|A(i) \cap S)|)) \geqslant \\ &\geqslant \sum_{i \in V} (\mathrm{H}(|A(i) \cap S| + 1) - \mathrm{H}(|A(i) \cap S)|)) = \sum_{i \in V} \frac{1}{|A(i) \cap S| + 1} \geqslant\end{aligned}$$

(z nierówności między średnią harmoniczną a arytmetyczną)

$$\begin{aligned}&\geqslant \frac{|V|^2}{\sum_{i \in V}(|A(i) \cap S| + 1)} = \frac{|V|^2}{\sum_{i \in V} |A(i) \cap S| + |V|} > \frac{|V|^2}{|V|(\ell - 1) + |V|} = \\ &= \frac{|V|}{\ell} \geqslant \frac{\ell \cdot n/k}{\ell} = n/k.\end{aligned}$$

Otrzymujemy zatem, że $\Delta(c_a) > \Delta(c_r)$. Ponadto gdy dodamy c_a do mniejszego komitetu, np. do $S \setminus \{c_r\}$, to wynik punktowy wzrośnie bardziej, niż gdybyśmy dodali c_a do S. Zatem gdy w komitecie S wymienimy c_r na c_a, to otrzymamy komitet o wyższym wyniku punktowym niż S. To daje sprzeczność, ponieważ reguła PAV w takim wypadku powinna wybrać $(S \setminus \{c_r\}) \cup \{c_a\}$ zamiast S. □

Literatura

[1] H. Aziz, M. Brill, V. Conitzer, E. Elkind, R. Freeman, and T. Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017.

[2] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. On the complexity of extended and proportional justified representation. In *Proceedings of the 32nd Conference on Artificial Intelligence (AAAI-2018)*, pages 902–909, 2018.