AGREGACJA SĄDÓW A AGREGACJA PREFERENCJI

Joanna Ochremiak*
Polska Akademia Nauk

Streszczenie: *Artykuł stanowi wprowadzenie do teorii agregacji sądów oraz porusza temat jej związków z teorią agregacji preferencji. Oparty na logice model formalny agregacji sądów porównany jest z modelem formalnym agregacji preferencji. Przedstawiony zostaje ponadto wynik w teorii agregacji sądów stanowiący dokładny odpowiednik twierdzenia Arrowa dla mocnych porządków.*

Słowa kluczowe: *agregacja sądów, agregacja preferencji, zachowanie jednomyślności, niezależność, twierdzenie Arrowa.*

JUDGMENT AGGREGATION AND PREFERENCE AGGREGATION

Abstract: *In the paper we present an introduction to the theory of judgment aggregation and discuss its relation to the theory of preference aggregation. We compare the formal model of judgment aggregation, based on logic, with the formal model of preference aggregation. Finally, we present a theorem in judgment aggregation which is an exact analogue of Arrow's theorem for strict preferences.*

Keywords: *judgment aggregation, preference aggregation, unanimity preservation, independence, Arrow's theorem.*

1. Wprowadzenie

Celem artykułu jest ukazanie związków pomiędzy stosunkowo nową dziedziną badań nad agregacją sądów a teorią agregacji preferencji.

W rozdziale pierwszym krótko przedstawiam historię oraz najważniejsze problemy obu teorii, sytuując je w szerszym kontekście teorii wyboru społecznego. W roz-

* Joanna Ochremiak, Instytut Matematyczny PAN, ul. Śniadeckich 8, 00-956 Warszawa, skr. poczt. 21, e-mail: joanna.ochremiak@gmail.com
Autorka bardzo dziękuje anonimowym Recenzentom za cenne uwagi.

dziale drugim koncentruję się na wprowadzeniu narzędzi potrzebnych do dalszej analizy. Zaczynam od przypomnienia podstawowych pojęć teorii agregacji preferecji, a następnie wprowadzam i omawiam model formalny teorii agregacji sądów wraz z przykładami. Rozdział trzeci zawiera porównanie obu modeli. W szczególności pokazuję, w jaki sposób można mówić o agregacji preferencji językiem teorii agregacji sądów. W rozdziale czwartym prezentuję twierdzenie, którego szczególny przypadek stanowi twierdzenie Arrowa dla mocnych porządków.

1.1. Agregacja preferencji

Teoria wyboru społecznego koncentruje się na poszukiwaniu odpowiedzi na pytanie, w jaki sposób należy podejmować zbiorowe decyzje na podstawie indywidualnych ocen społecznych rozwiązań. Jedną z jej najważniejszych gałęzi stanowi agregacja preferencji.

O agregacji preferencji mówimy, gdy poglądy członków grupy w kwestii rozważanych alternatyw społecznych wyrażone są za pomocą relacji binarnych. Każda osoba porządkuje zbiór możliwych rozwiązań, deklarując, które z nich uważa za lepsze, a które za gorsze. Mówiąc o preferencjach, zakładamy więc porządkowy pomiar intensywności preferencji. Innym sposobem reprezentacji indywidualnych ocen jest reprezentowanie ich za pomacą kardynalnych funkcji użyteczności, które stanowią mocniejszy, przedziałowy pomiar intensywności preferencji.

Skoro wiemy już, że podstawę społecznej decyzji stanowi profil indywidualnych relacji preferencji, należy zadać sobie pytanie, jaki ma być wynik procesu agregacji. Wyróżniamy dwa podejścia. Jeśli interesuje nas wyłącznie wybór z rozważanego zbioru alternatyw społecznych podzbioru najlepszych rozwiązań, to mówimy o *funkcjach wyboru*. Jeśli zaś efektem procesu podejmowania decyzji ma być relacja preferencji społecznej, czyli uporządkowanie zbioru alternatyw, to przedmiotem naszych badań czynimy *zasady zbiorowej oceny*, nazywane również po prostu *metodami agregacji*. Niniejszy artykuł oparty jest na drugim spośród opisanych powyżej alternatywnych sformułowań problemu decyzji. Zagadnienie agregacji indywidualnych preferencji w preferencję społeczną jest bowiem, jak się przekonamy, ściśle związane z teorią agregacji sądów. Warto jednak podkreślić, że chociaż oba podejścia wyznaczają różne dziedziny badań teoretycznych, to wyniki osiągane przy zastosowaniu pierwszego modelu można często przełożyć na drugi i na odwrót.

Za datę powstania teorii wyboru społecznego uznajemy rok 1951, kiedy wydana została monografia Arrowa *Social Choice and Individual Values*. Jest to jednocześnie data powstania teorii agregacji preferencji, gdyż to właśnie ta gałąź teorii wyboru społecznego została sformułowana jako pierwsza. Słynne twierdzenie Arrowa o niemożliwości odnosi się do zasad zbiorowej oceny przy dodatkowym założeniu indywidualnej i grupowej racjonalności.

Refleksja nad sposobami podejmowania zbiorowych decyzji ma jednak znacznie dłuższą historię. Prapoczątki teorii wyboru społecznego wiążemy z odkrytymi pod koniec XVIII w. paradoksami, ujawniającymi ułomności powszechnie stosowanych procedur decyzyjnych. Najbardziej bodaj znany przykład takiego paradoksu pochodzi od Condorceta:

Przykład 1.1. (Paradoks Condorceta) Trzy osoby podejmują decyzję w sprawie trzech alternatyw społecznych x, y oraz z. Ich preferencje są następujące:

- Osoba I uważa alternatywę x za najlepszą, zaś alternatywę y za lepszą niż z, wypisując w porządku od najlepszej do najgorszej: $x\ y\ z$.
- Osoba II uważa alternatywę y za najlepszą, zaś alternatywa z jest jej zdaniem lepsza niż x, wypisując w porządku od najlepszej do najgorszej: $y\ z\ x$.
- Osoba III uważa natomiast za najlepszą alternatywę z, zaś alternatywę x za lepszą niż y, wypisując w porządku od najlepszej do najgorszej: $z\ x\ y$.

Klasyczną metodą podejmowania decyzji zbiorowej jest uznanie za społecznie lepszą tej spośród pary alternatyw, którą więcej osób przedkłada nad tę drugą. Paradoksalnie jednak wyznaczona w ten sposób relacja preferencji okazuje się w tym przypadku cykliczna: rozwiązanie x jest społecznie lepsze od y, y społecznie lepsze od z, zaś z społecznie lepsze od x. Niemożliwy jest tym samym wybór jednej najlepszej alternatywy.

Własności konkretnych metod podejmowania zbiorowych decyzji rozważano, jak widać, na długo przed Arrowem. Przełomowe znaczenie jego monografii wynika z zupełnie nowego podejścia do analizy problemu. Podejście to nazywamy *metodą aksjomatyczną*. Jej zastosowanie wymaga sformułowania listy postulatów, które powinna z jakiegoś względu spełniać metoda agregacji. Warunki te mogą być rozmaite. Często, jak w przypadku twierdzenia Arrowa, odzwierciedlają niezbędne własności demokratycznego podejmowania decyzji. Mogą również wyrażać postulat odporności metody na strategiczne manipulacje, polegające na nieszczerym przedstawianiu własnych preferencji, czy też troskę o respektowanie praw indywidualnych. Dla danego zestawu warunków staramy się następnie wyróżnić metodę lub klasę metod je spełniających.

Może się oczywiście okazać, że nie istnieje żadna metoda o pożądanych własnościach. Wyniki badań przyjmują wówczas postać twierdzeń o nieistnieniu. Ich wymowa nie jest jednak wyłącznie negatywna. Wskazując na ograniczenia procedur decyzyjnych, zmuszają nas do urealnienia stawianych im wymagań. Wykazanie, że zadane warunki nie mogą zachodzić jednocześnie, skłania ponadto do głębszej refleksji nad problemem w celu ustalenia, które spośród wartości wyrażanych poprzez formalne postulaty są dla nas bardziej, a które mniej istotne.

Ogromna popularność Arrowa oraz wpływ, jaki on i jego następcy wywarli na teorię wyboru społecznego, sprawiły, że teoria agregacji preferencji na pół wieku zajęła dominującą pozycję w dziedzinie refleksji nad podejmowaniem decyzji zbiorowych.

1.2. Agregacja sądów

Nie każdy problem decyzji grupowej polega na wyznaczeniu relacji preferencji społecznej, która umożliwi wybór najlepszego spośród rozważanych rozwiązań. Często mamy do czynienia z sytuacją wymagającą po prostu wyrażenia wspólnej opinii w sprawie pewnego zbioru stwierdzeń. Podjęcie decyzji oznacza grupową zgodę bądź negację elementów zbioru kwestii do rozpatrzenia, nazywanego *agendą*.

Przykład 1.2. Rozważmy trzyosobowe grono ekspertów, którego zadaniem jest zajęcie wspólnego stanowiska w następujących kwestiach:

a : Nastąpi wzrost PKB.

$a \to b$: Jeśli nastąpi wzrost PKB, to wzrośnie inflacja.

b : Nastąpi wzrost inflacji.

Osoba I uważa, że nastąpi wzrost PKB, który spowoduje inflację. Osoba II sądzi co prawda, że wzrost PKB skutkuje inflacją, ale nie wierzy we wzrost inflacji i tym samym we wzrost PKB. Osoba III wierzy we wzrost PKB, ale nie zgadza się ze stwierdzeniem, że spowoduje on wzrost inflacji. Uważa, że inflacja nie wzrośnie. Przypuśćmy, że rozważane grono podejmuje decyzje w następujący, często spotykany sposób: zgadza się ze zdaniem, jeśli więcej członków grupy akceptuje to zdanie niż jego zaprzeczenie. Opinie ekspertów oraz ich ostateczne stanowisko w każdej z rozważanych kwestii ilustruje poniższa tabela.

Tabela 1. Dylemat dyskursywny

	a	$a \to b$	b
osoba I	tak	tak	tak
osoba II	nie	tak	nie
osoba IIII	tak	nie	nie
decyzja grupowa	tak	tak	nie

Źródło: Dietrich, F., C. List. 2007. Arrow's theorem in judgment aggregation. „Social Choice and Welfare" 29 (1): 19-33.

Dwie spośród trzech osób sądzą, że nastąpi wzrost PKB. Podobnie dwie osoby uważają, że wzrost PKB spowoduje wzrost inflacji oraz dwie osoby wierzą, że inflacja nie wzrośnie. Stanowisko wypracowane przez ekspertów jest zatem sprzeczne.

Powyższy przykład stanowi ilustrację szerszego problemu, nazwanego przez Pettita (zob. Pettit 2001) *dylematem dyskursywnym* (*discursive dilemma*). Okazuje się, że zastosowanie do zbioru logicznie powiązanych kwestii klasycznego sposobu podejmowania decyzji – *metody zwykłej większości*, może dawać w rezultacie sprzeczny zbiór zdań. Ponadto dzieje się tak w sytuacji, gdy indywidualne poglądy każdego członka grupy podejmującej decyzję są niesprzeczne. Historycznie pierwszy przykład dylematu dyskursywnego stanowi tak zwany *paradoks doktryny* (*doctrinal paradox*), analizowany w dziedzinie prawa i ekonomii (zob. Kornhauser, Sager 1986).

Przykład 1.3. (Paradoks doktryny) Trzech sędziów ma podjąć wspólną decyzję w sprawie winy podejrzanego. Muszą w tym celu zająć stanowisko w czterech kwestiach:

a : Podejrzany popełnił zarzucany mu czyn.

b : Zarzucany podejrzanemu czyn stanowi przestępstwo.

$c \leftrightarrow a \wedge b$: Podejrzany jest winny wtedy i tylko wtedy, gdy popełnił przestępstwo.

c : Podejrzany jest winny.

Wszyscy trzej sędziowie zgadzają się z trzecim stwierdzeniem stanowiącym doktrynę prawa. Sędzia I uważa, że podejrzany popełnił zarzucany mu czyn, który ponadto jest przestępstwem. Sędzia II sądzi, że podejrzany popełnił zarzucany mu czyn, ale jego zdaniem nie był to czyn niezgodny z prawem. Sędzia III twierdzi natomiast, że podejrzany nie popełnił zarzucanego mu czynu, ale zgadza się ze stwierdzeniem, że czyn ten stanowi przestępstwo. W konsekwencji sędzia I jest zdania, że podejrzany jest winny, zaś sędziowie II i III zaprzeczają winie podejrzanego.

Tabela 2. Paradoks doktryny

	a	b	$c \leftrightarrow a \wedge b$	c
sędzia I	tak	tak	tak	tak
sędzia II	tak	nie	tak	nie
sędzia IIII	nie	tak	tak	nie
decyzja grupowa	tak	tak	tak	nie

Źródło: Dietrich, F., C. List. 2007. Strategy-proof judgment aggregation. „Economics and Philosophy" 23: 269-300.

Załóżmy, że rozważane grono podejmuje decyzje, zgadzając się ze zdaniem, gdy więcej jego członków akceptuje to zdanie niż jego negację. Mamy wówczas do czynienia ze sprzecznością, którą ilustruje powyższa tabela. Sędziowie uznają, że podejrzany popełnił zarzucany mu czyn oraz że czyn ten stanowi przestępstwo, a jednocześnie twierdzą, że jest on niewinny, co stanowi sprzeczność z jednomyślnie akceptowaną doktryną prawa.

Paradoks doktryny sformułowany został początkowo w kontekście konfliktu między podejmowaniem decyzji na podstawie przesłanki a podejmowaniem decyzji, opierając się na wnioskach (zob. Kornhauser, Sager 1986). Dopiero po pewnym czasie zaobserwowano, że podobnie jak paradoks Condorceta ujawnia on istotną wadę powszechnie stosowanej procedury decyzyjnej (zob. Pettit 2001). Obserwacja ta stanowiła inspirację do głębszej refleksji na temat problemów decyzyjnych polegających na zajęciu grupowego stanowiska w sprawie zbioru logicznie powiązanych kwestii. Powstało pytanie, jak podejmować decyzje, by uniknąć sprzeczności?

Pierwsze formalne podejście do teorii agregacji sądów zaproponowali List i Pettit (zob. List, Pettit 2002). Podali oni ponadto dowód pierwszego, prostego twierdzenia o nieistnieniu. Oparty na rachunku zdań model, którym się posłużyli, był jednak bardzo ograniczony i pozwalał na analizowanie jedynie bardzo specyficznych, prostych zagadnień. Model, który zaprezentowany zostanie w ninejszym artykule, powszechnie stosowany współcześnie, stanowi jego uogólnienie. Został on sformułowany przez Dietricha (zob. Dietrich 2007a) i bazuje na niezwykle ogólnych podstawach logicznych. Dzięki temu w jego języku można wyrazić wiele skomplikowanych problemów decyzyjnych.

Teoria agregacji sądów doskonale wpisuje się w tradycję teorii wyboru społecznego nie tylko tematyką swoich badań, ale także stosowaną metodologią. W stosunku do procedur podejmowania decyzji, dotyczących zbioru logicznie powiązanych kwestii, formułowane są postulaty stanowiące bardziej lub mniej dokładne odpowiedniki warunków nakładanych na metody agregacji preferencji. Mają one odzwierciedlać wymagania demokracji, odporność na zachowania strategiczne czy troskę o respektowanie autonomii osób. Badania nad istnieniem i własnościami procedur decyzyjnych, spełniających określony zestaw postulatów, doprowadziły do udowodnienia wielu twierdzeń inspirowanych lub uogólniających znane wyniki teorii agregacji preferencji, jak twierdzenie Arrowa (zob. Dokow, Holzman 2010a i Dietrich, List 2007a), twierdzenie Gibbarda o istnieniu oligarchii (zob. Dietrich, List 2008a i Dokow, Holzman 2010b), paradoks liberalizmu Sena (zob. Dietrich, List 2008b) oraz twierdzenie Gibbarda-Satterthwaite'a o manipulowalności (zob. Dietrich, List 2007d).

Zastosowanie metody aksjomatycznej w dziedzinie agregacji sądów zaowocowało nowymi pytaniami badawczymi wynikającymi ze stopnia ogólności podstaw logicznych modelu. Odpowiedź na pytanie o klasę metod posiadających zadane własności zależy bowiem w tym przypadku nie tylko od samych własności, ale także od struktury logicznej problemu decyzyjnego, czyli stopnia powiązań logicznych zbioru rozważanych przez grupę kwestii. Możemy przykładowo zapytać, jak szeroka jest klasa agend, dla których jedyną metodą podejmowania decyzji spełniającą zmodyfikowane postulaty twierdzenia Arrowa jest dyktatura?

Przykład 1.4. Trzyosobowe grono specjalistów ma podjąć decyzję w trzech kwestiach:

a : Lekarstwo A jest skuteczne w walce z rakiem.

b : Lekarstwo B jest skuteczne w walce z rakiem.

c : Lekarstwo C jest skuteczne w walce z rakiem.

Specjalista I uważa, że skuteczne są wszystkie trzy leki. Specjalista II jest zdania, że skuteczny jest wyłącznie lek A, zaś specjalista III sądzi, że skuteczne są leki A oraz B. Przypuśćmy, że rozważane grono zgadza się ze zdaniem, jeśli więcej jego członków akceptuje to zdanie niż jego zaprzeczenie. Opinie specjalistów oraz ich ostateczne stanowisko w każdej z rozważanych kwestii ilustruje poniższa tabela.

Tabela 3. Metoda zwykłej większości dla prostego problemu decyzyjnego

	a	*b*	*c*
specjalista I	tak	tak	tak
specjalista II	tak	nie	nie
specjalista IIII	tak	tak	nie
decyzja grupowa	tak	tak	nie

Źródło: Opracowanie własne.

W przypadku problemu decyzyjnego opisanego powyżej zastosowanie metody zwykłej większości nie rodzi żadnych wątpliwości. Posiada ona wszystkie własności demokratycznego podejmowania decyzji, a struktura logiczna agendy jest tak prosta[1], że nie grozi nam podjęcie decyzji sprzecznej. Nietrudno się jednak domyślić, że im bardziej skomplikowany logicznie jest problem decyzyjny, tym mniejsza klasa procedur spełnia postawione wymagania. W ostateczności klasa ta może być pusta. Do najciekawszych wyników w teorii agregacji sądów należą twierdzenia jednoznacznie rozstrzygające, dla jakich agend istnieją, a dla jakich nie istnieją, metody o zadanych własnościach.

2. Podstawowe pojęcia

W tym rozdziale zaczniemy od przypomnienia podstawowych pojęć teorii agregacji preferencji. Następnie omówimy wprowadzony przez Dietricha (zob. Dietrich 2007a) model formalny teorii agregacji sądów. Większość definicji zostanie zilustrowana przykładami.

[1] Logicznie powiązane są tylko każde zdanie i jego negacja.

2.1. Agregacja preferencji – model formalny

Zbiór osób. Zbiór osób podejmujących społeczną decyzję oznaczamy przez $N = \{1, 2, ..., n\}$. Jest on skończony i niepusty.

Relacja preferencji. *Relacją preferencji* nazywamy dowolną relację binarną R na skończonym, niepustym zbiorze *alternatyw społecznych* $K = \{x, y, z, ...\}$. Przyjmujemy, że zapis xRy oznacza, że alternatywa x jest przynajmniej tak dobra, jak alternatywa y. Jeśli dla każdej spośród n osób podejmujących zbiorową decyzję mamy jej relację preferencji R_i na zbiorze alternatyw społecznych K, to otrzymujemy *profil preferencji indywidualnych* $(R_1, R_2, ..., R_n)$. Z relacją R wiążemy ponadto relację *mocnej preferencji* P, gdzie xPy zachodzi wtedy i tylko wtedy, gdy xRy i nieprawda, że yRx, a także relację *indyferencji* I taką, że xIy zachodzi wtedy i tylko wtedy, gdy xRy oraz yRx. Przyjmujemy, że zapis xPy oznacza, że alternatywa x jest lepsza niż y, zaś zapis xIy, że alternatywa x jest równie dobra jak y. Powiemy, że relacja preferencji R jest:

- *spójna*, jeśli dla dowolnych $x, y \in K$ takich, że $x \neq y$, zachodzi xRy lub yRx, czyli nie istnieją dwie różne alternatywy, których jednostka bądź grupa nie jest w stanie porównać,
- *zwrotna*, jeśli dla każdego $x \in K$ mamy xRx, czyli każda alternatywa jest uznawana za przynajmniej tak dobrą jak ona sama,
- *przechodnia*, jeśli dla dowolnych $x, y, z \in K$ zachodzi: (xRy oraz yRz) implikuje xRz,
- *quasi-przechodnia*, jeśli odpowiadająca jej relacja P jest przechodnia, czyli dla dowolnych $x, y, z \in K$ zachodzi: (xPy oraz yPz) implikuje xPz,
- *acykliczna*, jeśli odpowiadająca jej relacja P jest acykliczna, czyli dla dowolnych alternatyw $x, y, z, ..., u, v \in K$ zachodzi: jeśli (xPy, yPz, ..., uPv), to nieprawda, że vPz.

Nietrudno zauważyć, że relacja mocnej preferencji P jest *przeciwsymetryczna*, czyli dla dowolnych $x, y \in K$, jeśli xPy, to nieprawda, że yPx. Oznacza to, że jest ona również *przeciwzwrotna*[2], czyli dla każdego $x \in K$ nieprawda, że xPx. Obie własności zgadzają się z intuicją związaną z pojęciem mocnej preferencji. Pierwsza z nich oznacza, że uznanie alternatywy x za lepszą od y wyklucza uznanie alternatywy y za lepszą od x. Druga natomiast głosi, że żadna alternatywa nie jest uznawana za lepszą od siebie samej. Relacja indyferencji I jest z kolei *symetryczna*, czyli dla dowolnych $x, y \in K$, jeśli xIy, to zachodzi yIx. Uznając alternatywę x za równie dobrą jak y, uznajemy jednocześnie alternatywę y za równie dobrą jak x.

---

[2] Każda relacja przeciwsymetryczna jest przeciwzwrotna.

Zauważmy, że jeśli relacja preferencji R jest spójna i zwrotna, to dla dowolnej pary alternatyw $x, y \in K$ mamy xPy albo yPx albo xIy. Zamiast rozważać relację preferencji R możemy w tym przypadku równoważnie mówić o odpowiadającej jej relacji mocnej preferencji P. W dalszym ciągu artykułu zakładamy, że wszystkie analizowane relacje preferencji R są spójne i zwrotne, a mówiąc o relacji mocnej preferencji P, mamy na myśli relację przeciwsymetryczną. Relację R spełniającą ponadto warunek przechodniości nazywamy *słabym porządkiem* lub *racjonalną relacją preferencji*.

Metoda agregacji. Funkcję F, która profilowi preferencji indywidualnych $(R_1, R_2, ..., R_n)$ przyporządkowuje relację $F(R_1, R_2, ..., R_n) = R$ na zbiorze alternatyw społecznych K, interpretowaną jako relacja preferencji społecznej, nazywamy *metodą agregacji*. Zbiór profili, na którym określona jest funkcja F, nazywamy jej *dziedziną*.

Metoda zwykłej większości. Najpopularniejsza metoda agregacji preferencji – *metoda zwykłej większości* została już wspomniana przy okazji opisywania paradoksu Condorceta (zob. przykład 1.1). Teraz zdefiniujemy ją formalnie. Metoda ta przedkłada alternatywę x nad alternatywę y, jeśli więcej członków grupy przedkłada x nad y niż y nad x, czyli

$$xPy \Leftrightarrow |\{i \in N : xP_i y\}| > |\{i \in N : yP_i x\}|.$$

Dyktatura. Jeśli istnieje taka osoba i, że dla każdego profilu preferencji jej mocna indywidualna preferencja między dowolną parą alternatyw x, y wyznacza taką samą preferencję społeczną, czyli $xP_i y \Rightarrow xPy$, to mamy do czynienia z *dyktaturą*.

Pożądane jest, aby metoda agregacji preferencji spełniała pewne warunki wyrażające własności demokratycznego podejmowania decyzji.

- **Słaba optymalność Pareto.** Jeśli w sytuacji, gdy wszyscy przedkładają alternatywę x nad alternatywę y, metoda agregacji również przedkłada x nad y, to powiemy, że spełnia ona warunek *słabej optymalności Pareto*.
- **Niezależność od alternatyw niezwiązanych.** Metoda agregacji spełnia warunek *niezależności od alternatyw niezwiązanych*, jeśli dla dowolnych alternatyw $x, y \in K$ i dowolnych profili $(R_1, R_2, ..., R_n)$ oraz $(R_1^*, R_2^*, ..., R_n^*)$ należących do dziedziny zachodzi

$$(\forall i \in N (xR_i y \Leftrightarrow xR_i^* y \wedge yR_i x \Leftrightarrow yR_i^* x)) \Rightarrow (xRy \Leftrightarrow xR^* y \wedge yRx \Leftrightarrow yR^* x).$$

  Oznacza to, że społeczna preferencja między alternatywami x i y zależy wyłącznie od preferencji członków grupy względem tych alternatyw i nie mają na nią wpływu preferencje związane z innymi alternatywami.

- **Neutralność.** Metodę agregacji nazwiemy *neutralną*, jeśli dla każdej permutacji σ zbioru alternatyw K i dowolnych profili $(R_1, R_2, ..., R_n)$ oraz $(R_1^*, R_2^*, ..., R_n^*)$ należących do dziedziny zachodzi

$$(\forall i \in N \forall x, y \in K(xR_i y \Leftrightarrow \sigma(x)R_i^* \sigma(y))) \Rightarrow (\forall x, y \in K(xRy \Leftrightarrow \sigma(x)R^* \sigma(y))).$$

Oznacza to, że preferencja społeczna między każdą parą alternatyw wyznaczana jest w taki sam sposób.

2.2. Agregacja sądów – model formalny

Zbiór osób. Podobnie jak w przypadku agregacji preferencji, w teorii agregacji sądów rozważamy skończony, niepusty zbiór osób $N = \{1, 2, ..., n\}$ podejmujących społeczną decyzję.

Logika. Kwestie, które grupa ma do rozstrzygnięcia, wyrażone są w języku logiki formalnej. *Systemem logicznym* nazwiemy niepusty zbiór zdań[3] $\mathbf{L}$ zamknięty ze względu na negację (jeśli $p \in \mathbf{L}$, to $\neg p \in \mathbf{L}$) wraz z wyróżnioną rodziną swoich podzbiorów, nazywanych *niesprzecznymi*. Rodzina ta ma następujące własności:

- każda para $\{p, \neg p\} \subseteq \mathbf{L}$ jest zbiorem sprzecznym,
- podzbiór każdego zbioru niesprzecznego jest niesprzeczny,
- zbiór pusty $\varnothing \subseteq \mathbf{L}$ jest niesprzeczny oraz dla każdego niesprzecznego zbioru $S \subseteq \mathbf{L}$ istnieje niesprzeczny nadzbiór $T \subseteq \mathbf{L}$ taki, że dla każdej pary $\{p, \neg p\} \subseteq \mathbf{L}$ mamy $p \in T$ lub $\neg p \in T$.

O zbiorach $S \subseteq \mathbf{L}$ spoza rodziny zbiorów niesprzecznych mówimy natomiast, że są *sprzeczne*. Pojęcie niesprzeczności określa logiczne powiązania między zdaniami. Powiemy, że ze zbioru zdań $A \subseteq \mathbf{L}$ *wynika* zdanie $p \in \mathbf{L}$ (ozn. $A \perp p$), gdy zbiór $A \cup \{\neg p\}$ jest sprzeczny.

Przykład 2.1. Systemem logicznym jest rachunek zdań. Do zbioru $\mathbf{L}$ należą w tym przypadku wszystkie poprawnie zbudowane formuły rachunku zdań, czyli między innymi zdania a, $\neg a$, b, $a \wedge b$ oraz $a \rightarrow b$. Do rodziny zbiorów niesprzecznych zaliczamy wszystkie zbiory $S \subseteq \mathbf{L}$, dla których istnieje takie wartościowanie, że wartość logiczna każdego zdania $p \in S$ jest równa 1. Są to zbiory takie, jak $\{a, a \rightarrow b, b\}$ lub $\{a, a \wedge b\}$. Zbiorem sprzecznym będzie przykładowo zbiór $\{a, \neg a\}$, czy też $\{a, a \rightarrow b, \neg b\}$.

Zdefiniowane w powyższy sposób przez Dietricha (zob. Dietrich 2007a) pojęcie systemu logicznego jest niezwykle ogólne i abstrakcyjne. Kosztem wysokiego stopnia abstrakcji modelu otrzymujemy jednak dużą dowolność doboru odpowiedniego rodzaju logiki do analizy konkretnych problemów decyzyjnych. Poza rachunkiem zdań

[3] Przez zdania rozumiemy skończone ciągi symboli. Przykładowo: a, $x \rightarrow$ c, $\forall x(Px)$.

systemy logiczne w rozumieniu podanej definicji stanowią między innymi języki rachunku predykatów czy też logiki modalnej (zob. [Dietrich 2007a]).

Agenda. Zbiór kwestii, co do których grupa ma podjąć zbiorową decyzję, nazywamy *agendą*. Jest to skończony, niepusty podzbiór X zbioru **L** taki, że $X = \{p, \neg p: p \in X_+\}$ dla pewnego zbioru niezanegowanych[4] zdań X_+. Taka definicja pozwala uniknąć problemu podwójnych negacji. Dla uproszczenia, podając przykłady agend, będziemy zazwyczaj ograniczać się do wypisania elementów zbioru X_+. Zakładamy ponadto, że $\neg\neg p$ oznacza zdanie p.

Przykład 2.2. Do analizy problemu decyzyjnego opisanego w przykładzie 1.2 możemy zastosować system logiczny **L** będący rachunkiem zdań. Stwierdzenia rozważane przez grupę reprezentujemy za pomocą poprawnie zbudowanych formuł. Agendę stanowi wówczas zbiór

$$X = \{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}.$$

Elementami zbioru X_+ są natomiast zdania a, $a \to b$ oraz b.

Możliwe jest również analizowanie agend nieskończonych. Po pewnej modyfikacji założeń pozostaje dla nich w mocy większość twierdzeń. Uogólnienie wyników uzyskuje się jednak kosztem wydłużenia i komplikacji dowodów. Dla elegancji i jasności wywodu ograniczymy się do rozważania agend skończonych. Zakładamy ponadto, że do agendy nie należą zdania sprzeczne[5] ani tautologie[6], czyli dla każdego zdania $p \in X$ zarówno zbiór $\{p\}$, jak i $\{\neg p\}$ są niesprzeczne.

Zbiór sądów. *Zbiorem sądów* nazywamy zbiór zdań $A \subseteq X$ zawierający te spośród zdań należących do agendy, które są akceptowane przez pewną jednostkę bądź grupę. Przy czym zazwyczaj przyjmujemy, że „akceptować" zdanie $p \in X$ oznacza tyle, co „wierzyć, że zdanie p jest prawdziwe". Jeśli dla każdej spośród n osób podejmujących zbiorową decyzję mamy dany jej zbiór sądów, to otrzymujemy *profil* $(A_1, A_2, ..., A_n)$.

Przykład 2.3. W przypadku opisanym w przykładzie 1.2 zbiorami sądów osób I, II i III są odpowiednio zbiory $A_1 = \{a, a \to b, b\}$, $A_2 = \{\neg a, a \to b, \neg b\}$ oraz $A_3 = \{a, \neg(a \to b), \neg b\}$.

Za Dietrichem (zob. Dietrich 2007a) wprowadzimy teraz cztery klasyczne warunki nakładane na zbiór sądów. Stanowią one próbę formalnego odzwierciedlenia założeń o indywidualnej bądź grupowej racjonalności. Powiemy, że zbiór sądów $A \subseteq X$ jest:

4 Przez zdanie niezanegowane rozumiemy zdanie, które nie rozpoczyna się od symbolu negacji $\neg$.

5 Zdaniem sprzecznym nazywamy zdanie $p \in$ **L**, dla którego zbiór $\{p\}$ jest sprzeczny.

6 Tautologią nazywamy zdanie $p \in$ **L**, dla którego zbiór $\{\neg p\}$ jest sprzeczny.

- *niesprzeczny*, jeśli jest niesprzeczny w sensie danego systemu logicznego **L**,
- *zupełny*, jeśli dla każdego $p \in X$ mamy $p \in A$ lub $\neg p \in A$[7],
- *całkowicie racjonalny*, jeśli jest zupełny i niesprzeczny,
- *dedukcyjnie domknięty*, jeśli $A = \{p \in X: A \perp p\}$.

Zauważmy, że w rozumieniu powyższej definicji, aby dany zbiór sądów nazwać całkowicie racjonalnym, nie wystarczy, żeby był on niesprzeczny. Deklarująca go jednostka bądź grupa powinna wyrazić ponadto swoje zdanie na temat każdej kwestii należącej do agendy, gdyż zbiór sądów spełnia warunek zupełności, jeśli dla każdego zdania $p \in X$ akceptowane jest to zdanie lub jego negacja. Ciekawą alternatywę dla tego dość wymagającego i przez to często krytykowanego postulatu (zob. Gärdenfors 2006) stanowi dedukcyjna domkniętość. Oznacza ona, że akceptacja pewnego zbioru zdań pociąga za sobą jednocześnie akceptację wszystkich zdań, stanowiących jego logiczną konsekwencję. Nietrudno zauważyć, że każdy całkowicie racjonalny zbiór sądów jest jednocześnie dedukcyjnie domknięty.

Przykład 2.4. Powróćmy raz jeszcze do przypadku opisanego w przykładzie 1.2. Przypomnijmy, że $X = \{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}$. Wówczas zbiór sądów:

- $\{a, a \to b\}$ jest niesprzeczny, ale nie jest zupełny ani dedukcyjnie domknięty,
- $\{a, a \to b, \neg b\}$ jest zupełny, ale nie jest niesprzeczny ani dedukcyjnie domknięty,
- $\{a\}$ jest dedukcyjnie domknięty i niesprzeczny, ale nie jest zupełny,
- $\{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}$ jest zupełny i dedukcyjnie domknięty, ale sprzeczny (zbiór dedukcyjnie domknięty i sprzeczny zawiera wszystkie elementy agendy, czyli jedynym przykładem takiego zbioru jest cała agenda X),
- $\{a, \neg(a \to b), \neg b\}$ jest niesprzeczny i zupełny, czyli całkowicie racjonalny, a więc także dedukcyjnie domknięty.

Metoda agregacji. Funkcję F, która profilowi indywidualnych zbiorów sądów $(A_1, A_2, ..., A_n)$przyporządkowuje zbiór $F(A_1, A_2, ..., A_n) = A \subseteq X$, interpretowany jako zbiór zdań społecznie akceptowanych przez grupę N, nazywamy *metodą agregacji*. Chcielibyśmy, aby była ona maksymalnie uniwersalna, żeby osoby zaangażowane w proces decyzyjny mogły cieszyć się możliwie dużą swobodą w deklarowaniu swoich poglądów. Zakładamy jednak zazwyczaj, że indywidualne zbiory sądów będące przedmiotem agregacji są w jakimś sensie „rozsądne", czyli przynajmniej niesprzeczne. Zbiór profili, na którym określona jest funkcja F, nazywamy jej *dziedziną*.

7 Zauważmy, że własność zupełności zbioru sądów zależy w szczególności od agendy, której jest on podzbiorem. Wprowadzone pojęcie zupełności różni się więc nieco od występującej w logice własności zupełności, np. zupełności danego zbioru formuł rachunku zdań.

Metoda zwykłej większości. Metodą agregacji sądów bardzo często stosowaną w praktyce jest *metoda zwykłej większości*. Definiujemy ją analogicznie jak w przypadku agregacji preferencji. Zdanie $p \in X$ jest społecznie akceptowane, jeśli więcej członków grupy akceptuje to zdanie niż jego zaprzeczenie, czyli

$$F(A_1, A_2, \ldots, A_n) = \{p \in X : |\{i \in N : p \in A_i\}| > |\{i \in N : \neg p \in A_i\}|\}.$$

Jak już się przekonaliśmy, decyzja podjęta w wyniku zastosowania tej metody może być sprzeczna.

Metoda oparta na przesłankach. Jedną z metod, których dotyczy omawiany we wprowadzeniu paradoks doktryny, stanowi *metoda oparta na przesłankach* (*premise-based rule*). Aby ją zastosować, należy wyróżnić wśród rozważanych kwestii zbiór Y nazywany zbiorem *przesłanek* i podjąć w ich sprawie decyzję przy użyciu metody zwykłej większości. Ostatecznie społecznie akceptowane są wszystkie zdania, które stanowią logiczną konsekwencję zbioru zaakceptowanych przesłanek, czyli

$$F(A_1, A_2, \ldots, A_n) = \{p \in X : G(A_1, A_2, \ldots, A_n) \cap Y \perp p\},$$

gdzie funkcja G oznacza metodę zwykłej większości.

Prześledźmy zastosowanie tej metody w przypadku opisanym w przykładzie 1.2. Jeśli uznamy za przesłanki elementy zbioru $\{a, a \to b\}$ oraz ich negacje, to w wyniku zastosowania metody zwykłej większości otrzymamy zbiór $\{a, a \to b\}$. Jego logiczną konsekwencję stanowi zdanie b. Ostatecznie więc decyzją społeczną jest zbiór sądów $\{a, a \to b, b\}$ (zob. tabela 4).

W powyższym przykładzie jest stosunkowo jasne, które spośród elementów agendy należy uznać za przesłanki. Trzeba jednak podkreślić, że w ogólności podziału dokonuje się arbitralnie i zależy on od interpretacji. Jeśli powiązania logiczne pomiędzy elementami agendy są złożone, decyzja ta może się okazać trudna i kontrowersyjna. Zwłaszcza że w sposób oczywisty wpływa ona na wynik procesu agregacji.

Metoda oparta na konkluzjach. Metodę w pewnym sensie opozycyjną do metody opartej na przesłankach stanowi *metoda oparta na konkluzjach* (*conclusion-based rule*). Tu również wyróżniamy spośród rozważanych zdań pewien podzbiór Z nazywany zbiorem *konkluzji*, a następnie stosując metodę zwykłej większości decydujemy, które z nich będą społecznie akceptowane. Wstrzymujemy się natomiast od decyzji w kwestii pozostałych elementów agendy, interpretowanych jako przesłanki.

$$F(A_1, A_2, \ldots, A_n) = G(A_1, A_2, \ldots, A_n) \cap Z,$$

gdzie funkcja G oznacza metodę zwykłej większości.

Efektem zastosowania powyższej metody do przypadku opisanego w przykładzie 1.2, przy wyróżnionym zbiorze konkluzji $\{b, \neg b\}$, jest jednoelementowy zbiór sądów akceptowanych społecznie $\{\neg b\}$ (zob. tabela 4).

Do metody opartej na konkluzjach odnoszą się podobne zastrzeżenia, co do metody opartej na przesłankach. Wynik procesu agregacji zależy od wyboru zbioru wyróżnionych zdań Z. Problemy może rodzić także brak rozstrzygnięć w sprawie przesłanek. Grono podejmujące decyzję nie ma bowiem podstaw, by uzasadnić swoje stanowisko.

Dyktatura. Skrajnie zdegenerowaną metodę agregacji stanowi *dyktatura*. Mamy z nią do czynienia w sytuacji, gdy istnieje taka osoba i, że sąd $p \in X$ jest społecznie akceptowany wtedy i tylko wtedy, gdy akceptuje go osoba i, czyli dla każdego profilu należącego do dziedziny zachodzi $F(A_1, A_2, ..., A_n) = A_i$. Zauważmy, że pojęcie to definiujemy nieco inaczej niż w przypadku agregacji preferencji, gdzie dopiero mocna indywidualna preferencja dyktatora wyznacza taką samą preferencję społeczną. Relacja preferencji społecznej uzyskana w wyniku dyktatury nie musi być zatem dokładnie taka sama, jak relacja preferencji dyktatora.

Tabela 4 przedstawia wyniki zastosowania każdej z opisanych wyżej metod agregacji sądów w przypadku agendy oraz profilu indywidualnych zbiorów sądów opisanych w przykładzie 1.2.

Tabela 4. Porównanie wyników zastosowania czterech różnych metod agregacji sądów

	a	$a \rightarrow b$	c
osoba I	tak	tak	tak
osoba II	nie	tak	nie
osoba IIII	tak	nie	nie
metoda zwykłej wiekszości	tak	tak	nie
metoda oparta na przesłankach	tak	tak	tak
metoda oparta na konkluzjach			nie
dyktatura osoby II	nie	tak	nie

Źródło: Opracowanie własne.

Podobnie jak teoria agregacji preferencji, teoria agregacji sądów formułuje w stosunku do analizowanych metod formalne warunki mające na celu odzwierciedlenie pożądanych własności demokratycznego procesu podejmowania decyzji.

- **Zachowanie jednomyślności.** Jeśli zadeklarowanie przez każdego z członków grupy takiego samego zbioru sądów indywidualnych A wyznacza zbiór A jako decyzję społeczną, czyli $F(A, A, ..., A) = A$, to powiemy, że metoda agregacji *zachowuje jednomyślność*.
- **Niezależność.** Metodę agregacji nazwiemy *niezależną*, jeśli dla każdego zdania $p \in X$ i dowolnych profili $(A_1, A_2, ..., A_n)$ oraz $(A_1^*, A_2^*, ..., A_n^*)$ należących do dziedziny zachodzi

$$(\forall i \in N(p \in A_i \Leftrightarrow p \in A_i^*)) \Rightarrow (p \in F(A_1, A_2, ..., A_n) \Leftrightarrow p \in F(A_1^*, A_2^*, ..., A_n^*)).$$

Oznacza to, że społeczna decyzja w kwestii zdania $p \in X$ zależy wyłącznie od poglądów członków grupy na jego temat i nie mają na nią wpływu indywidualne oceny pozostałych zdań należących do agendy.

- **Systematyczność.** Metodę agregacji nazwiemy *systematyczną*, jeśli dla dowolnych zdań $p, q \in X$ i dowolnych profili $(A_1, A_2, ..., A_n)$ oraz $(A_1^*, A_2^*, ..., A_n^*)$ należących do dziedziny zachodzi

$$(\forall i \in N(p \in A_i \Leftrightarrow q \in A_i^*)) \Rightarrow (p \in F(A_1, A_2, ..., A_n) \Leftrightarrow q \in F(A_1^*, A_2^*, ..., A_n^*)).$$

Nietrudno zauważyć, że metoda systematyczna jest niezależna (wystarczy podstawić $p = q$). Ponadto decyzja społeczna w kwestii każdego zdania należącego do agendy podejmowana jest w taki sam sposób.

W literaturze spotkać można kilka, nieznacznie się od siebie różniących, wariantów warunków nakładanych na metody agregacji sądów. Powyższe definicje pochodzą z przeglądowej pracy Lista i Puppego (zob. List, Puppe 2009).

Dyktatura spełnia wszystkie powyższe postulaty. Metoda zwykłej większości również, ale tylko w przypadku, gdy jej dziedznę stanowią profile całkowicie racjonalnych zbiorów sądów. W przeciwnym razie narusza ona warunek niezależności (zob. przykład 3.6), a więc także systematyczności. Jeśli dziedziną metody opartej na przesłankach jest zbiór wszystkich profili całkowicie racjonalnych zbiorów sądów oraz wyznacza ona całkowicie racjonalne zbiory sądów, to zachowuje jednomyślność. Nie jest ona jednak ani niezależna (akceptacja poszczególnych zdań zależy od zbioru zaakceptowanych przesłanek), ani systematyczna (decyzja społeczna w sprawie przesłanek podejmowana jest inaczej niż w kwestii pozostałych elementów agendy). Stosując metodę opartą na konkluzjach, bierzemy pod uwagę tylko pewien podzbiór agendy. Metoda ta nie zachowuje zatem jednomyślności ani nie jest systematyczna. Podobnie jak metoda zwykłej większości bywa niezależna.

3. Porównanie modeli

Po zapoznaniu się z modelami formalnymi interesujących nas teorii możemy przejść do omówienia ich wzajemnych relacji. Wprowadzając ogólny model formalny teorii agregacji sądów, Dietrich zaproponował jednocześnie sposób ujęcia w nim modelu teorii agregacji preferencji (zob. Dietrich 2007a). W tym rozdziale pokażemy, w jaki sposób można wyrazić agregację preferencji, posługując się językiem teorii agregacji sądów. Porównamy ponadto warunki nakładane w obu teoriach na dziedzinę, wynik oraz metody agregacji.

3.1. Agregacja preferencji jako agregacja sądów

Agenda preferencji. Systemem logicznym, który służy do reprezentacji relacji preferencji, jest prosty język pierwszego rzędu z binarnym symbolem relacyjnym $\succ$, reprezentującym mocną preferencję oraz zbiorem stałych $K = \{x, y, z, ...\}$, reprezentujących alternatywy społeczne. Agendę preferencji stanowi zbiór $X = \{x \succ y, \neg(x \succ y) : x, y \in K \text{ i } x \neq y\}$.

Aby móc mówić o relacjach acyklicznych, quasi-przechodnich czy przechodnich, należy odpowiednio zdefiniować pojęcie niesprzeczności. Niech H będzie zbiorem zdań wyrażających w sposób formalny warunki spełniane przez ostre liniowe porządki:

- $(\forall v_1)(\forall v_2)(v_1 \succ v_2 \rightarrow \neg(v_2 \succ v_1))$ (przeciwsymetria),
- $(\forall v_1)(\forall v_2)(\forall v_3)((v_1 \succ v_2 \wedge v_2 \succ v_3) \rightarrow v_1 \succ v_3)$ (przechodniość),
- $(\forall v_1)(\forall v_2)(\neg(v_1 = v_2) \rightarrow (v_1 \succ v_2 \vee v_2 \succ v_1))$ (spójność).

Niech ponadto dla każdej pary różnych stałych $x, y \in K$ do zbioru H należy zdanie $\neg(x = y)$. Przyjmujemy, że zbiór zdań $A \subseteq X$ jest niesprzeczny wtedy i tylko wtedy, gdy zbiór $A \cup H$ jest niesprzeczny jako zbiór formuł rachunku predykatów pierwszego rzędu[8].

Każdej relacji preferencji R na zbiorze K możemy przyporządkować zbiór sądów A taki, że $A = \{x \succ y, \neg(y \succ x) : x, y \in K \text{ i } xPy\}$, gdzie P jest relacją mocnej preferencji, odpowiadającą relacji R.

Przykład 3.1. Rozważmy trzyelementowy zbiór alternatyw społecznych $K = \{x, y, z, ...\}$. Agendę preferencji stanowi wówczas zbiór X zawierający zdania $x \succ y$, $y \succ x$, $y \succ z$, $z \succ y$, $z \succ x$, $x \succ z$ oraz ich negacje. Relacji preferencji R takiej, że xPy, xPz oraz yIz przyporządkujemy zbiór sądów $\{x \succ y, \neg(y \succ x), x \succ z, \neg(z \succ x)\}$.

Zauważmy, że jeśli zachodzi xIy, to do zbioru sądów A nie należą ani zdania $x \succ y$, $\neg(x \succ y)$, ani zdania $y \succ x$, $\neg(y \succ x)$. Relacja indyferencji traktowana jest tym samym, jak wstrzymanie się od głosu w kwestii zdań $x \succ y$ oraz $y \succ x$.

Niesprzeczność a acykliczność preferencji. Acyklicznym relacjom preferencji R odpowiadają niesprzeczne zbiory sądów (zob. [Dietrich, List 2008b], [List, Polak 2010]).

Przykład 3.2. Rozważmy trzyelementowy zbiór alternatyw społecznych $K = \{x, y, z, ...\}$. Agendę preferencji stanowi wówczas zbiór X zawierający zdania $x \succ y$, $y \succ x$, $y \succ z$, $z \succ y$, $z \succ x$, $x \succ z$ oraz ich negacje.

8 Zbiór formuł rachunku predykatów pierwszego rzędu jest sprzeczny, jeśli za pomocą aksjomatów logicznych oraz reguł wnioskowania można z niego wyprowadzić formuły φ oraz $\neg\varphi$.

- Relacji preferencji R takiej, że xPy, yPz oraz zPx odpowiada zbiór sądów

$$\{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y), z \succ x, \neg(x \succ z)\}.$$

Relacja R nie jest acykliczna, a odpowiadający jej zbiór sądów jest sprzeczny, gdyż wchodzi w sprzeczność z warunkiem przechodniości należącym do zbioru H.

- Relacji preferencji R takiej, że xPy, yPz oraz xIz odpowiada zbiór sądów

$$\{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y)\}.$$

Relacja R jest acykliczna. Odpowiadający jej zbiór sądów jest niesprzeczny, gdyż stanowi podzbiór niesprzecznego zbioru sądów $\{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y), x \succ z, \neg(z \succ x)\}$, któremu z kolei odpowiada porządek liniowy, spełniający oczywiście wszystkie warunki należące do zbioru H.

Skoro potrafimy mówić o agregacji preferencji językiem teorii agregacji sądów i wiemy, że pojęciu acyklicznej relacji preferencji odpowiada pojęcie niesprzecznego zbioru sądów, to możemy zaobserwować, że opisany w przykładzie 1.1 paradoks Condorceta stanowi w istocie szczególny przypadek dylematu dyskursywnego (zob. List 2010, List, Polak 2010).

Przykład 3.3. Trzy osoby podejmują decyzję w sprawie trzech alternatyw społecznych x, y oraz z. Agendę preferencji stanowi zbiór X zawierający zdania $x \succ y$, $y \succ x$, $y \succ z$, $z \succ y$, $z \succ x$, $x \succ z$ oraz ich negacje. Preferencje członków grupy są następujące:

- Osoba I uważa alternatywę x za najlepszą, zaś alternatywę y za lepszą niż z, czyli $A_1 = \{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y), x \succ z, \neg(z \succ x)\}$.
- Osoba II uważa alternatywę y za najlepszą, zaś alternatywa z jest jej zdaniem lepsza niż x, czyli $A_2 = \{y \succ z, \neg(z \succ y), z \succ x, \neg(x \succ z), y \succ x, \neg(x \succ y)\}$.
- Osoba III uważa natomiast za najlepszą alternatywę z, zaś alternatywę x za lepszą niż y, czyli $A_3 = \{z \succ x, \neg(x \succ z), x \succ y, \neg(y \succ x), z \succ y, \neg(y \succ z)\}$.

Przypuśćmy, że rozważane grono podejmuje decyzje metodą zwykłej większości. Opinie poszczególnych osób oraz ostateczne stanowisko ilustruje poniższa tabela.

Tabela 5. Paradoks Condorceta

	$x \succ y$	$y \succ z$	$z \succ x$
osoba I	tak	tak	nie
osoba II	nie	tak	tak
osoba IIII	tak	nie	tak
grupa decyzyjna	tak	tak	tak

Źródło: List, C., B. Polak. 2010. Introduction to judgment aggregation. „Journal of Economic Theory” 145 (2): 441-466.

Zbiór sądów wyznaczony w wyniku zastosowania metody zwykłej większości to

$$A = \{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y), z \succ x, \neg(x \succ z)\}.$$

Jest on sprzeczny, gdyż wchodzi w sprzeczność z warunkiem przechodniości należącym do zbioru H. Odpowiadająca mu relacja preferencji jest cykliczna: rozwiązanie x jest społecznie lepsze od y, y – społecznie lepsze od z, zaś z – społecznie lepsze od x.

Dedukcyjna domkniętość a quasi-przechodniość preferencji. Zdefiniowanie niesprzeczności za pomocą zbioru H, ujmującego w sposób formalny warunki spełniane przez ostre liniowe porządki, umożliwia wyrażenie w języku teorii agregacji sądów pojęcia acyklicznej relacji preferencji. Jeśli relacja mocnej preferencji jest przechodnia, to odpowiadający jej niesprzeczny zbiór sądów jest ponadto dedukcyjnie domknięty. Zachodzi następujące stwierdzenie (zob. Dietrich, List 2008a):

Fakt 3.1. *Quasi-przechodnim relacjom preferencji R na zbiorze K odpowiadają niesprzeczne i dedukcyjnie domknięte zbiory sądów $A \subseteq X$. Odpowiedniość ta jest wzajemnie jednoznaczna.*

Przykład 3.4. Rozważmy trzyelementowy zbiór alternatyw społecznych $K = \{x, y, z, ...\}$. Agendę preferencji stanowi zbiór X zawierający zdania $x \succ y$, $y \succ x$, $y \succ z$, $z \succ y$, $z \succ x$, $x \succ z$ oraz ich negacje.

- Relacji preferencji R takiej, że xPy, yPz oraz xIz odpowiada zbiór sądów

  $$\{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y)\}.$$

  Relacja R nie jest quasi-przechodnia, a odpowiadający jej zbiór sądów nie jest dedukcyjnie domknięty, gdyż $\{x \succ y, y \succ z\} \cup H \perp x \succ z$, ale $x \succ z \notin \{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y)\}$.
- Relacji preferencji R takiej, że xPy, yIz oraz xIz odpowiada zbiór sądów $\{x \succ y, \neg(y \succ x)\}$. Relacja R jest quasi-przechodnia, a odpowiadający jej zbiór sądów jest dedukcyjnie domknięty.

3.2. Liniowe porządki

Dla oswojenia z agendą preferencji przyjrzymy się najprostszemu przypadkowi racjonalnej relacji preferencji, gdy odpowiadająca jej relacja mocnej preferencji P jest ostrym liniowym porządkiem. Dla dowolnej pary różnych alternatyw x, y zachodzi wówczas xPy albo yPx. Relację R o takiej własności nazywamy *mocnym porządkiem*.

Całkowita racjonalność sądów a mocne porządki. Niech X będzie agendą preferencji odpowiadającą zbiorowi alternatyw społecznych K. Rozważmy mocny porządek R na zbiorze K oraz odpowiadający mu zbiór sądów $A \subseteq X$. Zbiór A jest oczywiście niesprzeczny. Skoro dla każdej pary różnych alternatyw x, y zachodzi xPy albo yPx, to do zbioru A należy dokładnie jeden element z każdej pary zdań $\{x \succ y, \neg(x \succ y)\}$ oraz $\{y \succ x, \neg(y \succ x)\}$. Zbiór A jest więc również zupełny. Formalnie ujmuje to następujące stwierdzenie (zob. Dietrich, List 2007a):

Fakt 3.2. *Mocnym porządkom R na zbiorze K odpowiadają całkowicie racjonalne zbiory sądów $A \subseteq X$. Odpowiedniość ta jest wzajemnie jednoznaczna.*

Oznaczmy przez S zbiór wszystkich mocnych porządków na zbiorze K, a przez C zbiór wszystkich całkowicie racjonalnych zbiorów sądów $A \subseteq X$.

Przykład 3.5. Rozważmy trzyelementowy zbiór alternatyw społecznych $K = \{x, y, z, ...\}$. Agendę preferencji stanowi zbiór X zawierający zdania $x \succ y$, $y \succ x$, $y \succ z$, $z \succ y$, $z \succ x$, $x \succ z$ oraz ich negacje.

- Relacji preferencji R takiej, że xIy, yPz, xPz odpowiada zbiór sądów

$$\{y \succ z, \neg(z \succ y), x \succ z, \neg(z \succ x)\}.$$

  Relacja R nie jest mocnym porządkiem, a odpowiadający jej zbiór sądów nie jest całkowicie racjonalny, gdyż nie spełnia warunku zupełności.

- Relacji preferencji R takiej, że xPy, yPz oraz xPz odpowiada zbiór sądów

$$\{x \succ y, \neg(y \succ x), y \succ z, \neg(z \succ y), x \succ z, \neg(z \succ x)\}.$$

  Relacja R jest mocnym porządkiem, a odpowiadający jej zbiór sądów jest całkowicie racjonalny.

Własności metod agregacji. Niech F będzie metodą agregacji preferencji, określoną na zbiorze S^n wszystkich profili racjonalnych relacji preferencji, będących ponadto mocnymi porządkami, o wartościach w zbiorze S. Funkcji F w sposób naturalny odpowiada metoda agregacji sądów F' określona na zbiorze C^n wszystkich profili całkowicie racjonalnych zbiorów sądów, o wartościach w zbiorze C. Z faktu 3.2 wynika, że przyporządkowanie to jest bijekcją (zob. Dietrich, List 2007a, List 2010).

Zauważmy, że każdy profil całkowicie racjonalnych zbiorów sądów ma pewną własność ułatwiającą analizę. Skoro każda osoba akceptuje dokładne jedno spośród każdej pary zdań $\{p, \neg p\} \subseteq X$, to zbiór osób akceptujących zdanie p, czyli $\{i \in N : p \in A_i\}$ jednoznacznie wyznacza zbiór osób akceptujących zdanie $\neg p$. Mianowicie:

$$\{i \in N : \neg p \in A_i\} = N \setminus \{i \in N\ p \in A_i\}.$$

Wobec tego w przypadku mocnych porządków zachodzą następujące zależności pomiędzy znanymi nam warunkami nakładanymi na metody agregacji preferencji a warunkami nakładanymi na metody agregacji sądów (zob. Dietrich, List 2007a):

- Metoda agregacji preferencji F spełnia warunek niezależności od alternatyw niezwiązanych wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji sądów F' jest niezależna.
- Metoda agregacji preferencji F spełnia warunki słabej optymalności Pareto oraz niezależności od alternatyw niezwiązanych wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji sądów F' zachowuje jednomyślność i jest niezależna.
- Metoda agregacji preferencji F spełnia warunki niezależności od alternatyw niezwiązanych i neutralności wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji sądów F' jest systematyczna.

3.3. Częściowe porządki

Przypomnijmy, że quasi-przechodnim relacjom preferencji R odpowiadają przechodnie relacje mocnej preferencji P nazywane *ostrymi częściowymi porządkami*. Zbiór wszystkich quasi-przechodnich relacji preferencji na zbiorze K oznaczmy przez Q. Wspominaliśmy wcześniej (zob. fakt 3.1), że istnieje bijekcja pomiędzy zbiorem Q a zbiorem D wszystkich niesprzecznych i dedukcyjnie domkniętych zbiorów sądów $A \subseteq X$, gdzie X jest agendą preferecji odpowiadającą zbiorowi alternatyw społecznych K.

Własności metod agregacji. Niech F będzie metodą agregacji preferencji, określoną na zbiorze Q^n wszystkich profili quasi-przechodnich relacji preferencji, o wartościach w zbiorze Q. Funkcji F w sposób naturalny odpowiada metoda agregacji sądów F' określona na zbiorze D^n wszystkich profili niesprzecznych i dedukcyjnie domkniętych zbiorów sądów, o wartościach w zbiorze D. Z faktu 3.1 wynika, że przyporządkowanie to jest bijekcją.

W przypadku gdy indywidualne relacje mocnej preferencji P_i nie są porządkami liniowymi, a odpowiadające im zbiory sądów nie są zupełne, czyli dopuszczają wstrzymanie się od głosu, wzajemne relacje pomiędzy postulatami nakładanymi na oba rodzaje metod agregacji nieco się komplikują. Prześledźmy to na przykładzie warunku niezależności.

Metoda agregacji sądów F' spełnia warunek niezależności wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji preferencji F spełnia następujący warunek: dla dowolnych alternatyw $x, y \in K$ i dowolnych profili $(R_1, R_2, ..., R_n)$ oraz $(R_1^*, R_2^*, ..., R_n^*)$ należących do dziedziny Q^n zachodzi

$$(\forall i \in N(xP_i y \Leftrightarrow xP_i^* y)) \Rightarrow (xPy \Leftrightarrow xP^* y),$$

gdzie P_i jest relacją mocnej preferencji odpowiadającą relacji R_i.

Postulat równoważny warunkowi niezależności oznacza więc, że fakt przedkładania przez grupę alternatywy x nad alternatywę y powinien zależeć wyłącznie od tego, którzy spośród jej członków przedkładają alternatywę x nad alternatywę y. Nie ma mieć nań wpływu między innymi to, którzy spośród członków grupy mają preferencję odwrotną, czyli przedkładają alternatywę y nad alternatywę x. Warunek niezależności, który w przypadku mocnych porządków jest równoważny warunkowi niezależności od alternatyw niezwiązanych, okazuje się w ogólności mocniejszy.

Przykład 3.6. Przypuśćmy, że pięć osób podejmuje społeczną decyzję w sprawie trzyelementowego zbioru alternatyw społecznych $K = \{x, y, z, ...\}$. Rozważmy dwa profile preferencji $(R_1, R_2, ..., R_5)$ oraz $(R_1^*, R_2^*, ..., R_5^*)$. Załóżmy, że w pierwszym z nich dwie osoby przedkładają alternatywę x nad alternatywę y oraz trzy osoby przedkładają alternatywę y nad alternatywę x. Z kolei w drugim te same dwie osoby przedkładają alternatywę x nad alternatywę y, ale nikt nie przedkłada alternatywy y nad alternatywę x. Grupa podejmuje decyzję metodą zwykłej większości. Dla pierwszego profilu preferencji indywidualnych zachodzi wówczas yPx, zaś dla drugiego yP^*x. Jednocześnie w obu profilach dokładnie ten sam zbiór osób przedkłada alternatywę x nad alternatywę y. Metoda zwykłej większości spełnia warunek niezależności od alternatyw niezwiązanych, ale nie spełnia warunku niezależności.

Warunki niezależności i systematyczności w znanym nam brzmieniu zostały sformułowane z myślą o analizie metod agregacji sądów, których dziedzinę stanowią profile całkowicie racjonalnych, czyli w szczególności zupełnych, zbiorów sądów. W sytuacji, gdy istnieje możliwość wstrzymania się od głosu, konieczne staje się ich przeformułowanie. Rozsądne jest, aby decyzja społeczna w sprawie akceptacji zdania p zależała nie tylko od poglądów członków grupy na jego temat, ale także od ich oceny zdania $\neg p$:

- **Słaba niezależność.** Metodę agregacji nazwiemy *słabo niezależną*, jeśli dla każdego zdania $p \in X$ i dowolnych profili $(A_1, A_2, ..., A_n)$ oraz $(A_1^*, A_2^*, ..., A_n^*)$ należących do dziedziny zachodzi

$$(\forall i \in N(p \in A_i \Leftrightarrow p \in A_i^* \wedge \neg p \in A_i \Leftrightarrow \neg p \in A_i^*)) \Rightarrow$$
$$\Rightarrow (p \in F(A_1, A_2, ..., A_n) \Leftrightarrow p \in F(A_1^*, A_2^*, ..., A_n^*)).$$

  Oznacza to, że społeczna akceptacja zdania $p \in X$ zależy wyłącznie od poglądów członków grupy na temat zdań p oraz $\neg p$.

- **Słaba systematyczność.** Metodę agregacji nazwiemy *słabo systematyczną*, jeśli dla dowolnych zdań $p, q \in X$ i dowolnych profili $(A_1, A_2, ..., A_n)$ oraz $(A_1^*, A_2^*, ..., A_n^*)$ należących do dziedziny zachodzi

$$(\forall i \in N(p \in A_i \Leftrightarrow q \in A_i^* \wedge \neg p \in A_i \Leftrightarrow \neg q \in A_i^*)) \Rightarrow$$
$$\Rightarrow (p \in F(A_1, A_2, ..., A_n) \Leftrightarrow q \in F(A_1^*, A_2^*, ..., A_n^*)).$$

  Metoda słabo systematyczna jest słabo niezależna (wystarczy podstawić $p = q$). Ponadto decyzja społeczna w kwestii każdego zdania należącego do agendy podejmowana jest w taki sam sposób.

Jeśli dziedzinę metody agregacji sądów stanowi zbiór wszystkich profili całkowicie racjonalnych zbiorów sądów, to warunek słabej niezależności jest równoważny warunkowi niezależności, a warunek słabej systematyczności jest równoważny warunkowi systematyczności.

Warunki słabej niezależności i słabej systematyczności zdefiniowane zostały przez Dietricha i Lista (zob. Dietrich, List 2008a) w celu analizy metod agregacji, których dziedzinę stanowi zbiór D^n wszystkich profili niesprzecznych i dedukcyjnie domkniętych zbiorów sądów. Na nieco innym, choć równoważnym sformułowaniu powyższych postulatów, oparli swoje rozważania także Dokow i Holzman (zob. Dokow, Holzman 2010b). Autorzy ci wykazali w szczególności, że z jednego z udowodnionych przez nich twierdzeń wynika twierdzenie Gibbarda o istnieniu oligarchii dla teorii agregacji preferencji (zob. Dokow, Holzman 2010b).

W przypadku ogólnym zachodzą następujące zależności pomiędzy znanymi nam warunkami nakładanymi na metody agregacji preferencji, a sformułowanymi powyżej warunkami nakładanymi na metody agregacji sądów (Dokow, Holzman, 2010):

- Metoda agregacji preferencji *F* spełnia warunek niezależności od alternatyw niezwiązanych wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji sądów *F'* jest słabo niezależna.
- Metoda agregacji preferencji *F* spełnia warunki słabej optymalności Pareto oraz niezależności od alternatyw niezwiązanych wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji sądów *F'* zachowuje jednomyślność i jest słabo niezależna.
- Metoda agregacji preferencji *F* spełnia warunki niezależności od alternatyw niezwiązanych i neutralności wtedy i tylko wtedy, gdy odpowiadająca jej metoda agregacji sądów *F'* jest słabo systematyczna.

4. Twierdzenie Arrowa

W tym rozdziale ograniczymy nasze rozważania do całkowicie racjonalnych zbiorów sądów i relacji preferencji, stanowiących mocne porządki. Przyjrzymy się metodom agregacji sądów spełniającym postulaty odpowiadające warunkom nakładanym na metody agregacji preferencji w twierdzeniu Arrowa o niemożliwości. Przedstawimy twierdzenie jednoznacznie rozstrzygające, dla jakich agend jedyne metody spełniające ten zestaw postulatów są dyktatorskie, a dla jakich istnieją niezdegenerowane metody o takich własnościach. W tym rozdziale mówiąc o zbiorze osób N, zakładamy zawsze, że $n \geq 2$, a mówiąc o zbiorze alternatyw społecznych K, przyjmujemy, że $|K| \geq 3$. Zaczniemy od przypomnienia twierdzenia Arrowa dla metod agregacji preferencji określonych na zbiorze wszystkich profili mocnych porządków S^n, o wartościach w zbiorze S.

Twierdzenie 4.1. ***(Twierdzenie Arrowa)*** *Niech K będzie takim zbiorem alternatyw społecznych, że $|K| \geq 3$. Każda metoda agregacji preferencji F: $\mathrm{S}^n \to \mathrm{S}$ (określona na zbiorze wszystkich profili mocnych porządków) spełniająca warunki słabej optymalności Pareto i niezależności od alternatyw niezwiązanych jest dyktaturą.*

Przypomnijmy, że podana przez nas definicja dyktatury jako metody agregacji sądów jest mocniejsza, niż powszechnie znana definicja dyktatury jako metody agregacji preferencji. Jednak w przypadku mocnych porządków obie definicje są równoważne (zob. Dietrich, List 2007a).

Nasze zadanie polega na charakteryzacji agend, dla których każda, zachowująca jednomyślność i niezależna metoda agregacji sądów F: $\mathrm{C}^n \to \mathrm{C}$, określona na zbiorze wszystkich profili całkowicie racjonalnych zbiorów sądów, o wartościach w zbiorze całkowicie racjonalnych zbiorów sądów, jest dyktaturą.

Agenda totalnie zablokowana. Zbiór sądów $Y \subseteq X$ jest *minimalnym zbiorem sprzecznym*, jeśli jest sprzeczny oraz każdy jego podzbiór właściwy[9] jest niesprzeczny. Powiemy, że ze zdania $p \in X$ *warunkowo wynika* zdanie $q \in K$ (ozn. $p \perp^* q$), jeśli istnieje zbiór $Y \subseteq X$ taki, że zbiór $\{p\} \cup Y \cup \{\neg q\}$ jest minimalnym zbiorem sprzecznym oraz $p \neq \neg q$. Zauważmy, że zbiory $\{p\} \cup Y$ oraz $\{\neg q\} \cup Y$ są wówczas niesprzeczne oraz $\{p\} \cup Y \perp q$.

Relacja $\perp^*$ jest relacją binarną na zbiorze X. Agendę X nazywamy *totalnie zablokowaną* (*totally blocked*) wtedy i tylko wtedy, gdy dla dowolnych zdań $p, q \in X$ istnieje ciąg zdań $p_1, ..., p_k \in X$ takich, że $p = p_1 \perp^* p_2 \perp^* ... \perp^* p_k = q$ (ozn. $p \perp \perp^* q$). Własność totalnego zablokowania agendy zdefiniowana została przez Dietricha i Lista (zob.

9 Podzbiór Z zbioru Y jest właściwy, jeśli $Z \neq Y$.

Dietrich, List 2007a). Jej nieco inną wersję wprowadzili wcześniej Nehring i Puppe (zob. Nehring 2003). Dla rozważanego w niniejszym artykule przypadku agend skończonych oba podejścia są równoważne (zob. Dietrich, List 2007a).

Przykład 4.1. Agenda $X = \{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}$, opisana w przykładzie 1.2, nie jest totalnie zablokowana (List, Polak, 2010). Nie zachodzi między innymi relacja $\neg a \perp \perp^* a$.

Z naszego punktu widzenia kluczowe okaże się następujące stwierdzenie, którego dowód znaleźć można w pracy Nehringa (zob. Nehring 2003).

Fakt 4.1. *Jeśli* $|K| \geq 3$, *to agenda preferencji jest totalnie zablokowana.*

Parzysto-liczebna negowalność. Agenda X jest *parzysto-liczebnie negowalna* (*even-number negatable*), jeśli istnieje minimalny zbiór sprzeczny $Y \subseteq X$ taki, że dla pewnego podzbioru $Z \subseteq Y$ o parzystej liczbie elementów zbiór $(Y \setminus Z) \cup \{\neg p : p \in Z\}$ jest niesprzeczny (zob. Dietrich 2007a, Dietrich, List 2007a).

Przykład 4.2. Nietrudno udowodnić, że agenda $X = \{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}$, opisana w przykładzie 1.2, jest parzysto-liczebnie negowalna. Rozważmy minimalny zbiór sprzeczny $Y = \{\neg a, \neg(a \to b)\}$. Niech $p = \neg a$, a $q = \neg(a \to b)$. Wówczas zbiór $(Y \setminus \{p, q\}) \cup \{\neg p, \neg q\} = \{a, a \to b\}$ jest niesprzeczny.

Przyjrzyjmy się agendzie preferencji dla zbioru alternatyw społecznych K. Jeśli istnieją trzy różne alternatywy $x, y, z \in K$, to zbiór $Y = \{x \succ y, y \succ z, z \succ x\}$ jest minimalnym zbiorem sprzecznym. Niech $p = x \succ y$, ,zaś $q = z \succ x$. Wówczas zbiór

$$(Y \setminus \{p, q\}) \cup \{\neg p, \neg q\} = \{\neg(x \succ y), y \succ z, \neg(z \succ x)\}$$

jest niesprzeczny, gdyż jest podzbiorem zbioru $\{y \succ x, \neg(x \succ y), y \succ z, \neg(z \succ y), x \succ z, \neg(z \succ x)\}$, odpowiadającego acyklicznej relacji preferencji. Wynika stąd następujące stwierdzenie (zob. Dietrich, List 2007a):

Fakt 4.2. *Jeśli* $|K| \geq 3$, *to agenda preferencji jest parzysto-liczebnie negowalna.*

Dysponujemy w tym momencie wszystkimi narzędziami potrzebnymi do sformułowania zapowiadanego odpowiednika twierdzenia Arrowa.

Twierdzenie 4.2. *(i) Jeśli agenda X jest totalnie zablokowana oraz parzysto-liczebnie negowalna, to każda metoda agregacji sądów* F: $\mathrm{C}^n \to \mathrm{C}$ *(określona na zbiorze wszystkich* *profili całkowicie racjonalnych zbiorów sądów) zachowująca jednomyślność i niezależna jest dyktaturą.*

(ii) W przeciwnym razie (gdy agenda X jest nie jest totalnie zablokowana lub nie jest parzysto-liczebnie negowalna) dla $n \geq 3$ istnieją zachowujące jednomyślność i niezależne metody agregacji sądów F: $\mathrm{C}^n \to \mathrm{C}$ *(określone na zbiorze wszystkich profili całkowicie racjonalnych zbiorów sądów), niebędące dyktaturą.*

Ponieważ dla $|K| \geq 3$ agenda preferencji jest totalnie zablokowana i parzysto-liczebnie negowalna, to bezpośredni wniosek z powyższego twierdzenia stanowi twierdzenie Arrowa o niemożliwości dla mocnych porządków. Jego sformułowanie w języku teorii agregacji sądów brzmi następująco (zob. Dietrich, List 2007a, Dokow, Holzman 2010a):

Twierdzenie 4.3. ***(Twierdzenie Arrowa)*** *Niech X będzie agendą preferencji dla takiego zbioru alternatyw społecznych K, że* $|\mathrm{K}| \geq 3$. *Każda zachowująca jednomyślność i niezależna metoda agregacji sądów F:* $\mathrm{C}^n \to \mathrm{C}$ *(określona na zbiorze wszystkich profili całkowicie racjonalnych zbiorów sądów) jest dyktaturą.*

Jako pierwsi pełny dowód twierdzenia 4.2 podali Dokow i Holzman (zob. Dokow, Holzman 2010a). Niezależnie jego pierwszą część udowodnili Dietrich i List (zob. Dietrich, List 2007a). Oba rezultaty bazowały na wcześniejszych wynikach Nehringa i Puppego (zob. Nehring, Puppe 2010), od których pochodzi szczególny wariant twierdzenia, przy dodatkowym założeniu monotoniczności[10] metody agregacji (niepotrzebny staje się wówczas warunek parzysto-liczebnej negowalności). W teorii agregacji sądów znane są obecnie również twierdzenia, z których wywnioskować można orginalną, ogólną wersję twierdzenia Arrowa dla racjonalnych relacji preferencji (zob. Dietrich 2007b, Dokow, Holzman 2010b).

Jak szeroką klasę agend stanowią agendy totalnie zablokowane i parzysto-liczebnie negowalne? Warunek totalnego zablokowania jest stosunkowo wymagający. Jak widzieliśmy, nie spełnia go między innymi agenda opisana w przykładzie 1.2.

Przykład 4.3. Rozważmy agendę $X = \{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}$ opisaną w przykładzie 1.2. Wszystkie postulaty twierdzenia Arrowa spełnia metoda agregacji $F\colon \mathrm{C}^n \to \mathrm{C}$, dla której:

---

[10] Metodę agregacji nazwiemy *monotoniczną*, jeśli dla każdego zdania $p \in X$, dowolnych profili $(A_1, A_2, ..., A_n)$ oraz $(A_1^*, A_2^*, ..., A_n^*)$ należących do dziedziny i każdej osoby $i \in N$ zachodzi

$$(p \notin A_i \wedge p \in A_i^* \wedge p \in F(A_1, A_2, ..., A_n)) \Rightarrow p \in F(A_1^*, A_2^*, ..., A_n^*).$$

Oznacza to, że społecznie akceptowane zdanie $p \in X$ pozostaje społecznie akceptowane, gdy do grona osób je akceptujących dołączy kolejna osoba (zob. [List, Puppe 2009]).

- zdanie *a* jest akceptowane społecznie wtedy i tylko wtedy, gdy akceptują je wszyscy *członkowie grupy N (w przeciwnym razie społecznie akceptowane jest zdanie* $\neg a$*),*
- *zdanie* $\neg(a \to b)$ *jest akceptowane społecznie wtedy i tylko wtedy, gdy akceptują je wszyscy członkowie grupy N (w przeciwnym razie społecznie akceptowane jest zdanie* $a \to b$*),*
- *zdanie* $\neg b$ *jest akceptowane społecznie wtedy i tylko wtedy, gdy akceptują je wszyscy członkowie grupy N (w przeciwnym razie społecznie akceptowane jest zdanie b).*

Jednak nawet w przypadku agend, które nie są totalnie zablokowane, analiza relacji warunkowego wynikania wiele nam mówi o własnościach związanych z nimi metod agregacji. Przykładowo, jeśli pewien podzbiór *Y* agendy *X* jest totalnie zablokowany, to każda zachowująca jednomyślność i niezależna metoda agregacji sądów $F\colon \mathrm{C}^n \to \mathrm{C}$ dla agendy *X* jest systematyczna w odniesieniu do elementów zbioru *Y* (zob. Dietrich, List 2009, List, Polak 2010).

Warunek parzysto-liczebnej negowalności jest z kolei mało wymagający. Naruszają go wyłącznie agendy, w których powiązania logiczne między zdaniami są takie same, jak powiązania logiczne w pewnym zbiorze takich formuł klasycznego rachunku zdań, które zawierają wyłącznie symbole logiczne $\neg$ oraz $\leftrightarrow$. Formalne sformułowanie tej charakteryzacji wraz z dowodem znaleźć można w pracy Dokowa i Holzmana (zob. Dokow, Holzman 2010a).

Rozważmy agendę *X*, która nie jest parzysto-liczebnie negowalna. Z grupy podejmującej zbiorową decyzję wyróżnijmy podzbiór $M \subseteq N$, liczący nieparzyście wiele osób. Metodę agregacji sądów taką, że zdanie $p \in X$ jest społecznie akceptowane wtedy i tylko wtedy, gdy akceptuje je nieparzysta liczba osób ze zbioru *M*, nazywamy *regułą parzystości* (*parity rule*). Dla agend, które nie są parzysto-liczebnie negowalne, metody te jako jedyne spełniają warunki zachowania jednomyślności i niezależności. Mówi o tym następujące stwierdzenie, którego dowód znaleźć można w pracy Dokowa i Holzmana (zob. Dokow, Holzman 2010a):

Fakt 4.3. *Jeśli agenda X nie jest parzysto-liczebnie negowalna, to jedynymi metodami agregacji sądów* $F\colon \mathrm{C}^n \to \mathrm{C}$ *(określonymi na zbiorze wszystkich profili całkowicie racjonalnych zbiorów sądów) zachowującymi jednomyślność i niezależnymi są reguły parzystości.*

Obecnie znany jest cały szereg wyników modyfikujących lub uogólniających twierdzenie 4.2. Na metody agregacji nakładane są warunki takie, jak monotoniczność, systematyczność, anonimowość, brak weta czy neutralność w odniesieniu do zdania i jego negacji. Rozważa się agendy, dla których funkcje o zadanych własnościach nie są nie tylko dyktatorskie, ale także oligarchiczne. Przegląd współczesnych wyników dotyczących metod agregacji określonych na zbiorze profili całkowicie racjonalnych zbiorów sądów wraz z definicjami wszystkich wymienionych warunków znaleźć można między innymi w pracy Dietricha i Lista (2009). W teorii agregacji sądów analizowane są ponadto funkcje, których dziedzinę stanowi zbiór wszystkich profili niesprzecznych i dedukcyjnie domknietych (zob. Dietrich, List 2008a, Dokow, Holzman 2010b), a nawet tylko niesprzecznych (zob. Dietrich, List 2007c), zbiorów sadów.

5. Zakończenie

W ninejszym artykule starałam się przede wszystkim pokazać, w jaki sposób rozwijająca się teoria agregacji sądów opiera się na istniejącym fundamencie teorii agregacji preferencji. Świadomość paradoksalnych własności procedur podejmowania decyzji, polegających na uporządkowaniu zbioru alternatyw społecznych, umożliwiła dostrzeżenie paradoksów ogólniejszych. Postulaty formułowane w stosunku do metod agregacji preferencji znalazły swoje odzwierciedlenie w warunkach nakładanych na metody agregacji sądów, a twierdzenia dotyczące istnienia metod posiadających pożądane zestawy własności zainspirowały badaczy do stawiania pytań o możliwości ich uogólnienia. Wypracowano dzięki temu wiele wyników, które, podobnie jak twierdzenie zaprezentowane w niniejszym artykule, stanowią bardziej lub mniej dokładne odpowiedniki znanych faktów dotyczących agregacji preferencji. Co więcej, dowody uzyskanych w ten sposób twierdzeń często bazują na dowodach swoich pierwowzorów. Wymaga to jednak odnalezienia w znacznie przecież ogólniejszym i bardzo abstrakcyjnym modelu agregacji sądów odpowiedników pojęć, odnoszących się do relacji preferencji, co okazuje się zazwyczaj wysoce niebanalne.

W teorii agregacji sądów możemy więc widzieć pewien sposób uogólnienia teorii agregacji preferencji, czerpiący garściami z jej ponadpółwiekowej tradycji, ale pozwalający modelować szerszą klasę sytuacji społecznych, a także rodzący nowe pytania badawcze. Teoria agregacji sądów wiele zawdzięcza teorii agregacji preferencji. Czy jednak cokolwiek doń wnosi? Moim zdaniem tak. Teoria agregacji sądów dzięki swojemu bardzo ogólnemu, opartemu na logice, modelowi formalnemu pozwala przede wszystkim lepiej zrozumieć istotę problemów zwązanych z agregacją opinii jednostek, w tym również opinii wyrażających preferencje. Dostrzegając w relacji preferencji szczególny przypadek szerszego pojęcia zbioru sądów, możemy inaczej i głębiej spojrzeć na jej strukturę. Dzięki temu widzimy, jakie dokładnie własności preferencji nie pozwalają na podejmowanie w ich sprawie decyzji w demokratyczny sposób.

Ujęcie teorii agregacji preferencji w języku teorii agregacji sądów umożliwia ponadto pełniejszą analizę i uzasadnienie postulatów racjonalności nakładanych na relacje preferencji. Jeśli przyjmiemy, że istota słowa „racjonalność" związana jest z logiką, to rozsądne wydaje się powiązanie pytania o to, jakie sposoby uporządkowania alternatyw społecznych należy uznać za racjonalne, z wyborem określonego systemu logicznego, a w szczególności z definicjami pojęć sprzeczności i wynikania. Kontrowersje związane z warunkami acykliczności, przechodniości czy quasi-przechodniości relacji preferencji sprowadzamy do pytania o to, czy zaproponowany przez nas model logiki dobrze oddaje sposób myślenia jednostek, a także, jaki zbiór sądów należy uznać za racjonalny.

Bibliografia

Arrow, K. 1951. *Social Choice and Individual Values*. New York: Wiley.

Dietrich, F. 2007a. *A generalised model of judgment aggregation*. „Social Choice and Welfare" 28 (4): 529-565.

Dietrich, F. 2007b. *Aggregation theory and the relevance of some issues to others*. Working paper, University of Maastricht.

Dietrich, F., C. List. 2007a. *Arrow's theorem in judgment aggregation*. „Social Choice and Welfare" 29 (1): 19-33.

Dietrich, F., C. List. 2007b. *Majority voting on restricted domains*. Working paper, London School of Economics.

Dietrich, F., C. List. 2007c. *Judgment aggregation with consistency alone*. Working paper, London School of Economics.

Dietrich, F., C. List. 2007d. *Strategy-proof judgment aggregation*. „Economics and Philosophy" 23: 269-300.

Dietrich, F., C. List. 2008a. *Judgment aggregation without full rationality*. „Social Choice and Welfare" 31: 15-39.

Dietrich, F., C. List. 2008b. *A liberal paradox for judgment aggregarion*. „Social Choice and Welfare" 31: 59-78.

Dietrich, F., C. List. 2009. *Propositionwise judgment aggregation: the general case*. Working paper, London School of Economics.

Dokow, E., R. Holzman. 2010a. *Aggregation of binary evaluations*. „Journal of Economic Theory" 145 (2): 495-511.

Dokow, E., R. Holzman. 2010b. *Aggregation of binary evaluations with abstentions*. „Journal of Economic Theory" 145 (2): 544-561.

Gärdenfors, P. 2006. *An Arrow-like theorem for voting with logical consequences*. „Economics and Philosophy" 22 (2): 181-190.

Gibbard, A. 1969. *Social Choice and the Arrow Condition*. maszynopis niepublikowany.

Kornhauser, L.A., L.G. Sager. 1986. *Unpacking the Court*. „Yale Law Journal" 96 (1): 82-117.

Lissowski, G. 2001. *Problemy i metody teorii wyboru społecznego*. w: G. Lissowski (red.) *Elementy teorii wyboru społecznego*. Warszawa: Scholar.

List, C. 2010. *The theory of judgment aggregation: An introductory review*. Working paper, London School of Economics.

List, C., P. Pettit. 2002. *Aggregating Sets of Judgments: An Impossibility Result*. „Economics and Philosophy" 18 (1): 89-110.

List, C., B. Polak. 2010. *Introduction to judgment aggregation*. „Journal of Economic Theory" 145 (2): 441-466.

List, C., C. Puppe. 2009. *Judgment aggregaton: a survey*. w: P. Anand, C. Puppe and P. Pattanaik (red.) *Oxford Handbook of Rational and Social Choice*. Oxford: Oxford University Press.

Nehring, K. 2003. *Arrow's theorem as a corollary*. „Economics Letters" 80 (3): 379-382.

Nehring, K. 2006. *Oligarchies in Judgment Aggregation*. Working paper, University of California at Davis.

Nehring, K., C. Puppe. 2010. *Abstract Arrovian Aggregation*. „Journal of Economic Theory" 145 (2): 467-494.

Pattanaik, P.K. 1997. *Paradoksy agregacji preferencji*. w: G. Lissowski (red.) *Elementy teorii wyboru społecznego*. Warszawa: Scholar.

Pettit, P. 2001. *Deliberative Democracy and the Discursive Dilemma*. „Philosophical Issues" 11: 268-299.

Rubinstein, A., P. Fishburn. 1986. *Algebraic Aggregation Theory*. „Journal of Economic Theory" 38: 63-77.

Sen, A.K. 1970a. *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

Sen, A.K. 1970b. *The Impossibility of a Paretian Liberal*. „Journal of Political Economy" 78: 152-157.

Wilson, R. 1975. *On the Theory of Aggregation*. „Journal of Economic Theory" 10: 89-99.