Što dokazuje Arrowljev teorem

Zvonimir Šikić[1], Zagreb

Što je demokracija? Najopćenitije kazano, demokracija je način kolektivnog odlučivanja koji u što većoj mjeri poštuje individualne odluke. Kako doći do demokratskih odluka - nije lako pitanje. Za sada samo napominjemo kako se radi o tome da se iz individualnih odluka izvedu kolektivne, a da se pritom osiguraju jednaka prava svih individua. To se najčešće ne može postići u idealnom obliku, ali sustav je to demokratskiji što veći broj individua, pod što jednakijim uvjetima, sudjeluje u donošenju kolektivnih odluka.

Evo jedne konkretne ilustracije teškoća pri donošenju demokratskih odluka.

Pretpostavimo da neki kolektiv od 7 individua treba odlučiti o svojim preferencijama između opcija A, B, C, D i E, te da individualne preferencije izgledaju ovako:

$$1/7 \quad A > C > B > D > E$$

$$2/7 \quad A > D > C > E > B$$

$$1/7 \quad B > C > D > A > E$$

$$1/7 \quad B > D > C > A > E$$

$$1/7 \quad C > B > D > A > E$$

$$1/7 \quad E > B > D > C > A$$

Dakle, prva od sedam individua ponuđene opcije preferira redoslijedom $A > C > B > D > E$, tj. za nju je najbolja opcija A, nešto joj je lošija C, još lošija B, još lošija D i najlošija E. Sljedeće dvije individue preferiraju redoslijed $A > D > C > E > B$; sljedeća nakon njih preferira redoslijed $B > C > D > A > E$ itd. do zadnje individue koja preferira redoslijed $E > B > D > C > A$.

To su individualne preferencije. Problem je kako donijeti kolektivnu odluku o preferiranju ponuđenih opcija, tako da ona što više uzme u obzir postojeće individualne preferencije. Na primjer, koju bi opciju kolektiv trebao odabrati kao najbolju?

[1]Zvonimir Šikić, Fakultet strojarstva i brodogradnje Sveučilišta u Zagrebu

Odlučimo li se za većinski sustav donošenja odluka, najbolja je opcija A jer nju kao najbolju opciju preferira (relativna) većina individua. Naime, tri individue najboljom smatraju opciju A, dvije individue najboljom smatraju opciju B, a samo po jedna individua najboljima smatraju opcije C i E. Opciju D najboljom ne smatra nitko.

Odlučimo li se za eliminatorni sustav, najbolja je opcija B jer će u prvom krugu biti eliminirane opcije C, D i E, a tada će u drugom krugu pobijediti opcija B. Naime, u drugom će krugu prve tri individue kao najbolju preostalu opciju odabrati A, ali će zadnje četiri individue kao najbolju preostalu opciju (nakon eliminacije C, D i E) odabrati B.

Odlučimo li se za Bordin bodovni sustav, u kojemu prva opcija dobiva 4 boda, druga 3, treća 2, četvrta 1 i peta 0 bodova, najviše će bodova skupiti opcija C, pa je prema tom sustavu ona najbolja.

Odlučimo li se za niz duela (dvoboja) u kojima se prvo ogledaju opcije A i B, zatim se pobjednik tog duela ogleda s C, zatim pobjednik tog duela s D i na kraju pobjednik tog duela s E, ustanovit ćemo da na kraju pobjeđuje opcija D. Dakle, prema tom sustavu najbolja je opcija D.

Mogli bismo se odlučiti i za diktatorski sustav u kojemu je najbolja opcija ona koju preferira jedna istaknuta individua koju ćemo zvati diktatorom. Na primjer, ako je zadnja individua u našem početnom popisu individualnih preferencija diktator, onda će u diktatorskom sustavu pobijediti opcija E.

Sve u svemu, u različitim sustavima kolektivnog odlučivanja iz istih individualnih odluka izvode se različite kolektivne odluke:

Sustav	**Kolektivna odluka**
većinski	A
eliminatorni	B
bodovni	C
duelski	D
diktatorski	E

Koja od ovih kolektivnih odluka u najvećoj mjeri odražava individualne preferencije iz kojih je izvedena? Drugim riječima, koji od ovih sustava kolektivnog odlučivanja najviše uzima u obzir individualne preferencije na koje je primijenjen? Još kraće kazano, koji je sustav najdemokratskiji?

Na ova pitanja nije lako odgovoriti. **Kenneth J. Arrow** u svojoj je doktorskoj disertaciji iz 1950. godine dokazao da je matematički nemoguć sustav kolektivnog odlučivanja koji bi zadovoljavao razne međusobno suprotstavljene demokratske principe. Za taj i druge svoje rezultate 1972. godine dobio je Nobelovu nagradu za ekonomiju. To je bio prvi iz cijelog niza matematičkih rezultata kojima je dokazano da su mnogi na prvi pogled demokratski kriteriji međusobno suprotstavljeni, tj. da se ne mogu zajedno realizirati u istom sustavu kolektivnog odlučivanja.

To je iznimno važno znati i razumjeti jer je naivna potraga za takvim sustavima često vodila k tragičnim posljedicama. Na to je ukazao i K. J. Arrow u svojem govoru na dodjeli Nobelove nagrade 1972. godine:

…većina ljudi podcjenjuje nesigurnost svijeta. Velika zla proizašla su iz te vjere u sigurnost koja se pojavljuje kao vjera u povijesnu nužnost, velebne diplomatske sheme ili ekstremnu ekonomsku politiku.

Vratimo se ipak našem teškom pitanju. Koji od mnogih sustava kolektivnog odlučivanja najviše uzima u obzir individualne preferencije na koje se primjenjuje (većinski, eliminatorni, bodovni, ...)? Kraće kazano, koji je sustav najdemokratskiji?

Najprije moramo odrediti kriterije na temelju kojih ćemo uspoređivati različite sustave kolektivnog odlučivanja. Do danas su u političkoj i ekonomskoj praksi ponuđeni mnogi takvi kriteriji. Razmotrit ćemo one najpoznatije i široko prihvaćene.

To su ***Paretov*** i ***Condorcetov*** kriterij, te kriteriji ***monotonosti*** i ***relevantnosti***.

Paretov kriterij: Ako sve individue opciju X rangiraju ispred opcije Y, onda i kolektiv treba X rangirati ispred Y. Za sustav kolektivnog odlučivanja koji ispunjava ovaj zahtjev kaže se da je *Paretov*.

Condorcetov kriterij: Ako neka opcija u direktnom duelu sa svakom pojedinom opcijom izlazi kao pobjednik (jednostavnom većinom individualnih glasova), onda za nju kažemo da je *Condorcetov* pobjednik.

Važno je uočiti da raspodjela individualnih preferencija može biti takva da Condorcetov pobjednik ne postoji. Takva je, primjerice, sljedeća raspodjela individualnih preferencija

$$1/3 \quad A > B > C$$

$$1/3 \quad C > A > B$$

$$1/3 \quad B > C > A$$

Naime, A gubi u duelu s C, C gubi u duelu s B, a B gubi u duelu s A, pa nitko ne dobiva sve duele.

Za sustav u kojemu Condorcetov pobjednik, *ukoliko postoji*, uvijek pobjeđuje, kažemo da zadovoljava Condorcetov kriterij i zovemo ga Condorcetovim sustavom.

Kriterij monotonosti: Ako se individualne odluke promijene u korist opcije X, to ne smije pogoršati položaj opcije X u kolektivnoj odluci. Za sustav odlučivanja koji ispunjava ovaj zahtjev kaže se da je *monoton*.

Kriterij relevantnosti: Kako će kolektiv rangirati opcije X i Y ovisi samo o tome kako individue rangiraju te dvije opcije, a ne o tome kako one rangiraju ostale opcije. Za sustav kolektivnog odlučivanja koji ispunjava ovaj zahtjev kaže se da je *relevantan*. (Kriterij se češće naziva „neovisnost o irelevantnim alternativama".)

Navedeni se kriteriji čine toliko razumnima da većina ljudi smatra nedopustivim da sustav demokratskog odlučivanja ne ispunjava neki od tih kriterija. Međutim, nijedan od uobičajenih sustava ne ispunjava sve navedene kriterije. Pogledajmo kako stvari stoje s najpoznatijim sustavima:

	Paretov	Condorcetov	monoton	relevantan
većinski	DA	NE	DA	NE
eliminatorni	DA	NE	NE	NE
bodovni	DA	NE	DA	NE
duelski	NE	DA	DA	NE
diktatorski	DA	NE	DA	DA

Dakle, većinski sustav je Paretov, nije Condorcetov, monoton je, ali nije relevantan. Eliminatorni sustav je Paretov, ali ne zadovoljava nijedan od preostalih kriterija. I tako dalje. Svaka od činjenica iz gornje tablice može se matematički egzaktno dokazati, a ovdje ćemo, za primjer, dokazati samo dvije.

Dokažimo najprije da većinski sustav nije Condorcetov. U tu svrhu zamislimo 9 individua koje trebaju odlučiti između opcija A, B i C i pretpostavimo da su njihove individualne preferencije sljedeće:

$$4/9 \quad A > B > C$$

$$3/9 \quad B > C > A$$

$$2/9 \quad C > B > A$$

U većinskom sustavu pobijedit će opcija A. Lako je provjeriti da je Condorcetov pobjednik opcija B. Dakle, većinski sustav nije Condorcetov.

Dokažimo još da eliminatorni sustav nije monoton. U tu svrhu zamislimo sljedeću razdiobu individualnih preferencija:

$$40\% \quad A > B > C$$

$$29\% \quad B > A > C$$

$$31\% \quad C > B > A$$

Uz tu razdiobu, opcija B bit će eliminirana u prvom krugu izbora. Zatim će opcija A pobijediti u drugom krugu. Povećajmo sada broj individualnih lista koje opciju A stavljaju na prvo mjesto, na sljedeći način:

43% A > B > C

29% B > A > C

28% C > B > A

Iako se broj individualnih preferencija u korist opcije A povećao (s 40% na 43%), ona će u toj „povoljnijoj" situaciji izgubiti izbore jer će sada C otpasti u prvom krugu, pa će B pobijediti u drugom krugu.

Čini se da moramo smisliti nove sustave kolektivnog odlučivanja težimo li uistinu demokratskim sustavima koji ispunjavaju sve postavljene kriterije. No, to nije moguće. Kenneth J. Arrow dokazao je kako je matematički nemoguć sustav kolektivnog odlučivanja koji bi bio Paretov, relevantan i nediktatorski.

Kao što smo već rekli, to je bio prvi iz cijelog niza matematičkih rezultata kojima je dokazano da su mnogi demokratski kriteriji međusobno suprotstavljeni, tj. da se ne mogu zajedno realizirati u istom sustavu kolektivnog odlučivanja. Do toga nužno dolazi čim su u igri više od dvije opcije.

Ako se treba odlučiti između samo dvije opcije, onda se može matematički egzaktno dokazati da je većinski sustav kolektivnog odlučivanja jedini sustav koji jamči monotonost, jednakost svih individualnih odluka i neutralnost prema obje opcije (u tu se svrhu jednakost i neutralnost mogu egzaktno definirati, čime se ovdje nećemo opterećivati). Taj je dokaz prvi objavio K. May 1952. godine.

Pri odabiru između tri ili više opcija, idealnih rješenja (koja teže ispunjavanju svih demokratski prihvatljivih kriterija), čini se, nema.

Razmotrimo malo detaljnije što se sve krije iza ove negativne tvrdnje.

Arrowljev teorem dokazuje da nema izbornog sustava koji je **Paretov** (ako svi glasači opciju A stavljaju ispred opcije B, onda i rezultat izbora mora biti takav da je A ispred B), **relevantan** (kako su na izborima rangirane opcije A i B ovisi samo o tome kako glasači rangiraju te dvije opcije, a ne o tome kako rangiraju preostale opcije) i **nije diktatorski** (sustav je diktatorski ako se izabrani redoslijed opcija poklapa s redoslijedom koji preferira jedan jedini „glasač", diktator, potpuno neovisno o redoslijedima koje preferiraju ostali glasači).

No, osim toga, teorem pretpostavlja da su preferencije glasača ordinalne, a ne kardinalne. To znači da oni ponuđene opcije (npr. A, B i C) rangiraju po redoslijedu (npr. A > B > C), a ne tako da im pridaju numeričke vrijednosti (npr. A = 9 bodova, B = 3 boda i C = 0 bodova).

Teorem također pretpostavlja da su preferencije svih glasača tranzitivne, te da je redoslijed opcija koje sustav izvodi iz tih preferencija također tranzitivan (dakle iz A > B i B > C u svakom slučaju mora slijediti A > C).

Tek uz ovu detaljniju analizu možemo reći u kojem smislu Arowljev teorem dokazuje da su demokratski izbori nemogući (kako se popularno i netočno često govori).

Dakle, oni su nemogući:

1. ako smatramo da se individualne preferencije mogu iskazivati samo ordinalno, a ne i kardinalno;
2. ako smatramo da su „Pareto", „relevantnost" i „ne-diktatorstvo" nužni uvjeti demokratičnosti izbora.

Prvi stav proizlazi iz teškoća koje su začetnici teorije odlučivanja imali s uspoređivanjem kardinalno izraženih preferencija različitih osoba (Kako usporediti mojih 5 bodova s tvoja 4 boda?). Arrow je zato *a priori* odbacio kardinalno iskazivanje preferencija *jer ono nema smisla*. Doista neobično, s obzirom da je Von Neumann godinama prije Arrowljevog rezultata pokazao kako se bodovne preferencije različitih osoba mogu uspoređivati (u *Theory of Games and Economic Behavior* iz 1944.).

Osim toga, gotovo je trivijalno naći opcije za koje su kardinalne preferencije smislenije od ordinalnih. Na primjer, to su sljedeće opcije:

A. Dobit ćete 100 kuna.

B. Dobit ćete 100 kuna i bit ćete premlaćeni do smrti.

C. Bit ćete premlaćeni do smrti.

Što bolje iskazuje vaše preferencije, ordinalni $A > B > C$ ili kardinalni $A = 10^{10}$ bodova; $B = 1$ bod i $C = 0$ bodova?

Dakle, izborne sustave s kardinalnim (bodovnim) iskazivanjem preferencija sigurno ne treba *a priori* odbaciti. No, ako njih uzmemo u obzir, onda Arrowljev teorem više nije bitan.

Na primjer, izborni sustav u kojem glasači svakoj opciji slobodno dodjeljuju određeni broj bodova (npr. od 0 do 9 bodova) i koji opcije konačno rangira prema broju dobivenih bodova jest „Paretov", „relevantan" i „ne-diktatorski". U tom smislu za njega ne vrijedi Arrowljev teorem. (Naravno, Arrowljev teorem *a priori* isključuje takve sustave.)

Drugi je stav također upitan. Točnije, „Pareto" i „ne-diktatorstvo" očito su neupitna demokratska načela, ali „relevantnost" to nije. Pokazat ćemo to na povijesno važnom primjeru kojim je Condorcet 1780-ih (vidjet ćemo, neopravdano) diskreditirao Bordin izborni sustav.

Radi se o primjeru sa sljedećom razdiobom glasova:

30	A > B > C
1	A > C > B
10	C > A > B
1	C > B > A
10	B > C > A
29	B > A > C

Bordin izborni sustav (koji 1. poziciji daje 2 boda, 2. daje 1 bod i 3. daje 0 bodova) daje konačni redoslijed B > A > C. Može se dokazati da bilo koja (od prve do treće pozicije opadajuća) razdioba bodova daje isti rezultat.

Condorcetov prigovor je da u direktnim duelima A pobjeđuje B (jer 41 glasač stavlja A ispred B, a 40 B ispred A) i A pobjeđuje C (jer 60 glasača stavlja A ispred C, a samo 12 C ispred A). Dakle, A je Condorcetov pobjednik (u direktnim duelima pobjeđuje sve druge opcije), a Bordin ga sustav ne proglašava pobjednikom. Condorcet zaključuje da Bordin sustav zato ne valja.

Uočite da se tu radi upravo o „relevantnosti". Condorcetov pobjednik (u njegovom primjeru to je A) određuje se međusobnim uspoređivanjem dviju opcija, neovisno o tome kako su rangirane ostale opcije. Točno to zahtijeva „relevantnost". Drugim riječima, Condorcetov prigovor jest prigovor da Bordin sustav ne poštuje Arrowljev (gotovo dva stoljeća mlađi) uvjet „relevantnosti".

Je li Condorcet u pravu, tj. je li „relevantnost" relevantna? Donald Saari je početkom našega stoljeća, u okviru svoje geometrijske teorije socijalnog izbora, dokazao da nije.

Razmislite najprije o običnom većinskom izboru između dvije opcije A i B. Pretpostavimo da je u igri 3000 glasača. Ako među njima detektiramo 1000 onih koji preferiraju A>B i 1000 onih koji preferiraju B>A, tih 2000 glasova možemo zanemariti (jer jedni druge poništavaju). Rezultat izbora određuje preostala tisuća.

Potpuno analogno, 30 glasača s preferencijama

10	A > B > C
10	B > C > A
10	C > A > B

jedni druge poništavaju, pa i njih možemo zanemariti. Naime, svaka od opcija A, B i C ima isti broj (u ovom slučaju 10) prvih, drugih i trećih mjesta.

Jednako tako možemo zanemariti i 3 glasača s preferencijama

1	$A > C > B$
1	$C > B > A$
1	$B > A > C$

jer sada svaka od opcija ima po jedno 1., 2. i 3. mjesto.

No, ako u Condorcetovom primjeru zanemarimo te glasače jer im se glasovi poništavaju, odluku donose preostali glasači:

20	$A > B > C$
28	$B > A > C$

Sada je očito da je C najlošija opcija i da je $A>B$ u svakom razumnom sustavu.

Ukratko, Bordin rezultat $A > B > C$ je valjan, a Condorcetov „relevantni" $B > A > C$ nije.

Saari je u svojoj geometrijskoj teoriji dokazao da to vrijedi sasvim općenito:

Ako se iz skupa individualnih preferencija uklone sve one koje se međusobno poništavaju (jer imaju jednaki broj 1., 2., 3. itd. mjesta za sve ponuđene opcije), onda Bordin sustav primijenjen na preostale individualne preferencije zadovoljava sve Arrowljeve uvjete.

Spomenimo na kraju da je Condorcet-Arrowljeva „relevantnost", pored svega dosad kazanog, još i neusklađena s uvjetom tranzitivnosti. Naime, duele možete odigrati (i Codorcetovog pobjednika naći) čak i ako preferencije glasača nisu tranzitivne. Saari je, u vezi s tim, dokazao sljedeće:

Ako se uvjet „relevantnosti" modificira tako da, osim rangiranja parova, uzima u obzir i tranzitivnost individualnih preferencija, onda Bordin sustav zadovoljava sve Arrowljeve uvjete.

Sve u svemu, Arowljev je teorem daleko od dokaza nemogućnosti demokratskog odlučivanja.