Il teorema dell'impossibilità di Arrow

Mario Tirelli

Aprile, 2010

1 Nozioni preliminari

Gli individui/votanti sono in numero $I > 2$, e $\mathcal{I}$ indica l'insieme degli individui, $\mathcal{I} = \{1, .., i, ...I\}$. Ciascun individuo i ha un ordinamento di preferenza riflessivo transitivo e completo $\succeq_i$ sull'insieme delle alternative possibili, X, $\succeq := (\succeq_i)_i$. Le alternative sulle quali sono chiamati a votare sono 3 o più di 3, $\#X \geq 3$, e comprende le tre alternative A, B, C. L'ordinamento di preferenza sociale è indicato con $\succeq_S$, anch'esso definito su X. L'insieme di possibili ordinamenti di preferenza individuali $\succeq$ (riflessive e transitive) è indicato con $\mathcal{R}^I = \Pi_i \mathcal{R}$, con $\mathcal{P}^I = \Pi_i \mathcal{P}$ quelli stretti $\succ$. $\mathcal{A} \subseteq \mathcal{R}^I$ denota l'insieme ammissibile di preferenze individuali $\succeq$.

Definizione 1 *Una funzione del benessere sociale è una funzione $F : \mathcal{A} \to \mathcal{R}$ che associa a ogni profilo ammissibile di preferenze individuali $\succeq$ in $\mathcal{A}$, una preferenza sociale $\succeq_S = F(\succeq)$ in $\mathcal{R}$.*

Definizione 2 *(Dominio universale) F ha dominio universale se la classe di preferenze individuali ammissibili $\mathcal{A} = \mathcal{R}^I$, cioè comprende qualsiasi ordinamento $\succeq$ definito su X.*

Un esempio di dominio universale con $X = \{A, B, C\}$ è quello che si ha quando, per ciascun individuo i, $\mathcal{P}$ comprende i seguenti ordinamenti,
$A \succ_i B \succ_i C$
$A \succ_i C \succ_i B$
$B \succ_i A \succ_i C$
$B \succ_i C \succ_i A$
$C \succ_i A \succ_i B$
$C \succ_i B \succ_i A$
Vi sono 6 ordinamenti per ogni individuo. Quindi, ad esempio, con 2 individui il dominio $\mathcal{P}^2$ comprende 36 ordinamenti individuali. Il dominio cresce poi esponenzialemente nel numero di individui.

Definizione 3 *(Paretiano) F genera preferenze sociali $\succeq_S$ che rispettano il principio Paretiano, se ad ogni coppia di alternative x, y in X, $x \succeq_i y$ per tutti gli i implica $x \succeq_S y$.*

Definizione 4 *(Indipendenza delle Alternative Irrilevanti –iai) F soddisfa iai se genera preferenze sociali che consentono di ordinare ciascuna coppia di alternative x, y in X solo in base all'ordinamento di preferenza individuale per tale coppia x, y. In altri termini, per qualsiasi coppia di possibili ordinamenti individuali $\succeq, \succeq'$ in $\mathcal{R}^I$ che non differiscono per l'ordinamento che gli individui danno a x, y (per ogni i, $x \succeq_i y \Leftrightarrow x \succeq'_i y$ oppure $y \succeq_i x \Leftrightarrow y \succeq'_i x$), si ha lo stesso ordinamento sociale $\succeq_S$ in $\mathcal{R}$ ($x \succeq_S y \Leftrightarrow x \succeq'_S y$ oppure $y \succeq_S x \Leftrightarrow y \succeq'_S x$).*

Definizione 5 *(Dittatoriale) F si dice dittatoriale se esiste un individuo, $\hat{\imath}$, tale che la preferenza sociale generata da $\succeq$, $\succeq_S = F(\succeq)$, è uguale a $\succeq_{\hat{\imath}}$ per qualsiasi profilo iniziale $\succeq$ in $\mathcal{A}$.*

2 Teorema dell'impossibilità di K. Arrow

Teorema 6 *(Teorema Impossibilità di K. Arrow (1951)) Supponiamo di avere almeno tre alternative. Non esiste alcuna funzione del benessere sociale F che sia non-dittatoriale e che soddisfi le seguenti proprietà (o assiomi):*
1) dominio universale
2) paretiano (o dell'unanimità)
3) transitività e completezza
4) indipendenza delle alternative irrilevanti (iai).

Dimostrazione[1]:

Gli individui/votanti sono $i = 1, ..., I$, e i loro ordinamenti di preferenza individuali sono transitivi e completi e tutti ammissibili ($\succeq$ in $\mathcal{R}^I$, e $\succ$ in $\mathcal{P}^I$). Le alternative sulle quali sono chiamati a votare sono 3 o più di 3; comunque comprendono $A, B, C : X \supseteq \{A, B, C\}$.

Dividiamo la dimostrazione in tre parti.

<u>Parte 1</u>

Partiamo col mostrare che **se un ordinamento di preferenze degli individui mette B in posizione estrema (top o bottom dell'ordinamento) allora l'ordinamento sociale metterà B in una posizione estrema.** Per contraddizione, supponiamo non sia così (l'ordinamento soc. non mette B in posizione estrema); ad esempio, $A \succ_S B, B \succ_S C \Rightarrow_{transitività} A \succ_S C$. Per l'ipotesi *iai*, tale ordinamento sociale tra (A, B) non varia se spostiamo C sopra A nell'ordinamento di preferenza di ogni individuo. Tuttavia, facendo così C sarà unanimamente preferita ad A e per la proprietà paretiana, $C \succ_S A$. Cio' contraddice la transitività della regola sociale. Quindi, non è possibile che la regola sociale, $\succ_S$, non abbia B in posizione estrema quando ciò avviene per i votanti.

[1] John Geanakoplos, "Three Brief Proofs of Arrows Impossibility Theorem," *Economic Theory*, (2005), 26(1): 211-215. [CFDP 1123RRR and CFP 1116], http://cowles.econ.yale.edu/~gean/publications.htm

Un individuo/votante si dice **pivotal** se cambiando il suo ordinamento di preferenza (e quindi il suo voto) cambia anche l'ordinamento sociale: più precisamente, $i^*(x)$ è pivotal su x se per qualsiasi alternativa y in X, $x \succ_{i^*(x)} y$ implica $x \succ_S y$. Supponiamo che $i^*(B)$ sia l'individuo pivotal su B. Tale individuo esiste sempre per ogni alternativa $(A, B, C, ..)$. Infatti, consideriamo un ordinamento che metta B nella posizione in assoluto peggiore; l'ordinamento sociale -all'unanimità- stabilisce che $x \succ_S B$ per qualsiasi alternativa $x \neq B$. Costruiamo due profili di preferenze individuali, I, II. Spostando B da tale posizione bottom ad una top, progressivamente, per ciascun individuo, a partire da $i = 1$; si arriva ad un individuo in corrispondenza del quale tale variazione comporta una variazione dell'ordinamento sociale; tale indivio è il pivot $i^*(B)$ e nella peggiore delle ipotesi (essendo la regola sociale paretiana), tale indivio coinciderà con l'ultimo, I.
I - E' costruito a partire da un ordinamento nel quale tutti gli individui hanno B come alternativa peggiore, facendo spostare B agli individui $1, 2, ..i^*(B) - 1$, al top del proprio ordinamento. Il risultato della Parte 1 implica che B è in posizione estrema nella preferenza sociale.

$$\begin{array}{ccccccc}
1 & \cdots & i^*-1 & \boxed{i^*} & i^*+1 & \cdots & I \\
B & & B & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & B & B & & B
\end{array}$$

II - E' uguale a I eccetto che ora anche $i^*(B)$ ha spostato B al top del suo ordinamento.
Ovviamente $i^*(B)$ è decisivo su (A, B)**:** ovvero, $A \succ_{i^*(B)} B \Rightarrow A \succ_S B$. Inoltre è decisivo anche rispetto a qualsiasi altra coppia alternativa contenente B, (x, B); dato che nel passaggio tra I e II il ranking sociale di B cambia perché $i^*(B)$ è pivotal rispetto a B (in II B assume una posizione al top dell'ordinamento sociale ed è quindi preferita a qualsiasi alternativa x; B rimane preferita rispetto a qualsiasi x anche se si permutano tra loro mozioni (x, x') , diverse da B, negli ordinamenti individuali, per l'assioma iai).

Parte 2

Mostriamo ora che $i^*(B)$ **è pivotal su qualsiasi coppia di alternative che non include** B, ad esempio, (A, C). A tal fine, costruiamo un terzo profilo di preferenze,
III - III è uguale a II eccetto che ora $i^*(B)$ ha spostato A sopra a B, e che tutti gli altri individui modificano il loro ordinamento relativo di (A, C), lasciando B

in posizione estrema. L'ordinamento in III è del tipo,

$$
\begin{array}{ccccccc}
1 & \cdots & i^*-1 & \boxed{i^*} & i^*+1 & \cdots & I \\
B & & B & A & \vdots & & \vdots \\
\vdots & & \vdots & B & \vdots & & \vdots \\
\downarrow C & & & & \downarrow A & & \downarrow A \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
A \uparrow & & & C & C \uparrow & & C \uparrow \\
\vdots & & \vdots & \vdots & B & & B
\end{array}
$$

Dato che per tutti gli $i \neq i^*(B)$, B è in posizione estrema essi possono invertire la loro preferenza individuale per (A, C) senza che cambi quella per (B, C) e (A, B).

Osservazioni:

a) In I : B era in posizione estrema e quindi estrema per la società, e $x \succ_{i^*(B)} B$ per tutte le alternative x (dato che $i^*(B)$ è decisivo su $(A, B), (B, C)$ e su qualsiasi alternative contenente B). Ne segue che $A \succ_S B, C \succ_S B$

b) In II : $B \succ_S A, B \succ_S C$, sempre a causa del fatto che $i^*(B)$ è decisivo su (x, B), per qualsiasi x

c) III e I sono uguali a meno della posizione relativa assunta da A rispetto a C (e alle altre alternative); quindi, per iai, anche per profilo III deve valere $\boxed{A \succ_S B}$

d) III e II sono uguali a meno della posizione relativa assunta da A rispetto a C (e alle altre alternative) per tutti gli $i \neq i^*(B)$ e nel profilo di $i^*(B)$ per la posizione relativa assunta da A rispetto a B e C (e alle altre alternative); mentre non cambia la posizione di B rispetto a C. Quindi, per iai l'ordinamento sociale (B, C) dovrà essere lo stesso in II e in III. Dato che in II $i^*(B)$ era decisivo su qualsiasi (x, B) e preferiva B a qualsiasi altra alternativa, $\boxed{B \succ_S C}$

e) Usando i punti $c, d)$ e la transitività della $(\succ_S)$, $\boxed{A \succ_S C}$

Quindi, esistono ordinamenti di preferenze individuali rispetto ai quali $i^*(B)$ è decisivo anche su (A, C) : in III $A \succ_{i^*(B)} C$ implica che $A \succ_S C$. Dato che la scelta di (A, C) è arbitraria (potevo specificare e scegliere altre coppie di alternative che non includono B), $i^*(B)$ è decisivo su qualsiasi coppia di alternative.

Per riassumere, notate che l'ordinamento sociale indotto dagli ordinamenti di preferenza individuale in III rispetto a A, B, C, stabilisce che $A \succ_S B, B \succ_S C, A \succ_S C$. Ciò segue dal fatto che: $i)$ quando ordino (A, B), III è analogo a I per l'assioma iai; $ii)$ quando ordino (B, C), III è analogo a II per l'assioma iai; $iii)$ l'ordinamento di (A, C) segue dalla transitività della preferenza sociale.

<u>Parte 3</u>

Per mostrare l'esistenza di un **dittatore**, i^*, mi basta stabilire che esiste un solo decisore che risulta pivotal su tutte le alternative; ovvero che $i^*(x) = i^*(B)$ per qualsiasi alternativa x in X. Proviamo con $x = C$ e cerchiamo di dimostrare che $i^*(C) = i^*(B)$. A tal fine, ripercorriamo lo stesso argomento della *Parte 2*, partendo da un ordinamento che mette C in posizione estrema e dimostriamo che $i^*(C)$ è decisivo su tutte le alternative che includono C e poi su tutte le altre. Ciò implica che $i^*(C)$ è decisivo anche su (A, B): $A \succ_S B$ se $A \succ_{i^*(C)} B$. Ma, come abbiamo visto, esistono dei profili di preferenze per i quali $i^*(B)$ è decisivo su (A, B), quindi i due individui devono coincidere (il dominio della preferenza sociale è non ristretto e deve quindi includere ordinamenti come I,II,III). Infatti, prendiamo un profilo nel quale tutti mettono C in posizione estrema, rispetto al quale, per definizione $i^*(C)$ è decisivo rispetto a qualsiasi alternativa, anche (A, B). Modifichiamolo come segue: per tutti gli i con $A \succ_i B$ sposto B al bottom, mentre per tutti gli i' con $B \succ_{i'} A$ sposto B al top. Abbiamo costruito un profilo con B in posizione estrema rispetto al quale $i^*(B)$ è pivotal senza modificare la posizione relativa di A rispetto a B per alcuno (ovvero $A \succ_{i^*(C)} B$). Quindi $A \succ_S B$ se $A \succ_{i^*(B)} B$, ovvero $i^*(B) = i^*(C)$. Dato che C era stata scelta arbitrariamente, esiste un *dittatore* che impone alla società il proprio ordinamento di preferenza. ■

3 Commenti

Cosa stabilisce il teorema di Arrow e cosa si impara dalla sua dimostrazione? Il teorema stabilisce che l'unico ordinamento di preferenza sociale possibile, avente gli assiomi sopra elencati $1) - 4)$, è di tipo dittatoriale; ovvero, non esiste un sistema "democratico" che rispecchi tali proprietà di scelta sociale (la regola di votazione a maggioranza, non fa ovviamente eccezione). Tuttavia il teorema non consente di stabilire l'identità del dittatore. Il pivot varia al variare di come gli individui sono ordinati in $\mathcal{I}$, o chiamati ad esprimersi. La funzione del benessere sociale considerata può benissimo essere anonima e quindi l'ordinamento sociale non risentire dell'ordinamento degli individui rispetto a $\mathcal{I}$; tuttavia, permutando l'ordine degli individui cambia l'identità del pivot (questo appare chiaramente dalla dimostrazione). Il teorema stabilisce quindi solo l'esistenza di un dittatore; tutti gli individui sono potenzialmente tali.

L'argomento usato nella dimostrazione si basa sostanzialmente sulla costruzione di tre profili di preferenza, del tipo I, II, III. A partire da questi tre profili–possibile per l'assioma di *dominio universale*—viene costruita una preferenza sociale. Inoltre usa più volte gli assiomi di *transitività* e di *indipendenza delle alternative irrilevanti*. Ma cosa accade se cambiamo gli assiomi e cosa comporta concettualmente tale cambiamento? E' possibile ristabilire l'esistenza di una regola democratica e a che prezzo? Gran parte della letteratura successiva il contributo di Arrow ha cercato di rispondere a queste domande, proprio con l'obiettivo di superare il risultato dell' "impossibilità".

Deroghe agli assiomi di scelta sociale: il *dominio universale* Una prima deroga riguarda l'assioma di *dominio universale*. L'ipotesi maggiormente esplorata è quella di preferenze individuali che siano *single-peaked*, cioè che abbiano un punto di massima soddisfazione (assoluta) rispetto a un insieme chiuso di alternative, dotato di un ordinamento lineare (riflessivo, transitivo, totale o completo). L'esito raggiunto con il teorema dell'elettore mediano (Duncan Black) è incoraggiante in quanto consente di mostrare che votazioni a *maggioranza semplice* consentono di ottenere un ordinamento sociale completo e transitivo delle molteplici alternative possibili. Ciò avviene perché l'elettore mediano, cioè l'elettore il cui picco di preferenza è mediano rispetto a quelli degli altri elettori, riesce a raccogliere la maggioranza di consensi sull'alternativa preferita, per qualsiasi coppia di alternative; si dice che l'elettore mediano è un *Condorcet-winner:* dalla sua posizione di centro (mediana) riesce ad orientare l'esito delle votazioni. Cerchiamo di chiarire meglio questi concetti.

Ad esempio, Sia $X = [0, 100]$ l'insieme di alternative possibili; supponete che x in X misuri la dimensione in MgW di una centrale elettrica sulla quale il comitato $\mathcal{I}$ viene chiamato ad esprimersi. Questo insieme ha una proprietà importante, possiede un *ordine lineare*: i suoi elementi possono essere ordinati secondo un ordinamento, $\geq$, che è riflessivo, transitivo, completo. Nell'esempio considerato X è un sottoinsieme della retta dei numeri reali che, ovviamente, possiede tale ordinamento.

Definizione 7 *Un ordine di preferenza $\succeq$ si dice a picco unico (single-peaked) rispetto a un ordinamento lineare $(\leq)$ su X se esiste un'alternativa $\hat{x}$ in X con la proprietà che ciascuna coppia di alternative z, y in X soddisfa la seguente proprietà $z \succ y$ se e solo se vale una delle due situazioni seguenti,*

$$\begin{gathered} y < z \leq \hat{x} \\ \hat{x} \leq z < y \end{gathered}$$

In questa definizione, $\hat{x}$ è l'alternativa rispetto alla quale il "picco" di preferenza è definito; nel caso $X = \mathbb{R}^N$, $z \succ y$ se z è ordinabile come più vicino, in termini di distanza Euclidea, al picco $\hat{x}$ rispetto a y. Più semplicemente, nell'esempio precedente, tutte le alternative nell'intervallo $[0, 100]$ alla sinistra di $\hat{x}$ sono ordinabili in base a in senso di preferenza crescente da sinistra a destra (secondo l'ordinamento $\geq$ su X); le alternative alla destra di $\hat{x}$ sono ordinabili in senso decrescente di preferenza verso destra (cioè, secondo l'ordinamento $\geq$ su X). Possibili andamenti delle preferenze sono di tipo campanulare con picco di massimo, oppure monotone crescenti o decrescenti, rispettivamente con picco $\hat{x} = 0$, e $\hat{x} = 100$.

Una proprietà interessante è poi la seguente. Sia $\succeq$ definita su $X \subset \mathbb{R}$; $\succeq$ è single-peaked se e solo se è strettamente convessa. Tuttavia, quando le alternative consistono in interi menu (o vettori di mozioni), $X \subset \mathbb{R}^{N>1}$, le cose cambiano e non vi è più una corrispondenza stretta tra convessità e la proprietà di *single-peaked.*[2]

Definizione 8 *(Elettore mediano) L'elettore mediano, m, è quello il cui picco di preferenza $\hat{x}^m$ si trova in posizione centrale rispetto ai picchi degli altri individui,*

[2]Si veda la discussione in Mas-Colell, Whinston e Green, "Microeconomic Theory", p.804.

$\{\hat{x}^i\}_{i\in\mathcal{I}}$, *ordinati secondo* $\geq$:

$$m : \#\{i \in \mathcal{I} : \hat{x}^i \geq \hat{x}^m\} \geq \frac{I}{2} \quad e \quad \#\{i \in \mathcal{I} : \hat{x}^m \geq \hat{x}^i\} \geq \frac{I}{2}$$

E' facile verificare che un elettore mediano è sempre definito; se I è dispari è unico e si tratta di quello con il picco $(\frac{I-1}{2}+1)-esimo$.

Proposition 9 *In un contesto nel quale gli elettori hanno preferenze single-peaked e si vota a maggioranza assoluta, il picco dell'elettore mediano non può essere battuto da alcun altra alternativa (si dice che l'elettore mediano è il Condorcet winner).*

Proof. Considera qualsiasi alternativa y in X, e supponi che $\hat{x}^m > y$. Dobbiamo mostrare che y non può affermarsi a maggioranza su $\hat{x}^m$, ovvero che $\#\{i \in \mathcal{I} : y \succ \hat{x}^m\} \leq \#\{i \in \mathcal{I} : \hat{x}^m \succ y\}$. Osservate che tutti gli individui con picco alla destra di $\hat{x}^m$, preferiscono $\hat{x}^m$ a y. Questi sono la metà del totale, cioè $I/2$; ai quali si aggiunge l'elettore mediano. La maggioranza assoluta a favore di $\hat{x}^m$ è quindi costituita. Lo stesso argomento può essere usato nel caso di qualsiasi alternativa z in X, $z > \hat{x}^m$. ■

Una seconda proprietà interessante delle preferenze single-peaked è che il Paradosso di Condorcet sulla regola di voto a maggioranza, secondo una procedura di voto completa, non può verificarsi per tali preferenze. Questo perché la votazione a maggioranza (preferenza sociale) in base a preferenze single-peaked ha sempre un elemento massimamente preferito (individuato da $\hat{x}^m$), ovvero che non può essere battuto da alcun altra alternativa votata a maggioranza.

Questo risultato, riabilita in parte il criterio a maggioranza rispetto al teorema di Arrow. Con preferenze individuali single-peaked il voto a maggioranza assoluta determina una regola di aggregazione del consenso (preferenza sociale) che è transitiva e completa.

Altre deroghe agli assiomi di scelta sociale Gli assiomi dell'*unanimità (paretiano)*, dell'*indipendenza delle alternative irrilevanti (iai)* e della *transitività* della preferenza sociale. Tutti questi assiomi hanno un motivo di essere.

▲ L'importanza del primo è evidente, anche se è opportuno notare che esso può essere indebolito rispetto alla versione ipotizzata da Arrow (nella versione usata da Arrow esso prevede che se ciascun individuo preferisce strettamente x a y allora la società preferisce strettamente x a y). Sen indebolisce tale assioma introducendo "la norma del consenso": $y \succ_S x$ salvo che x non sia debolmente Pareto superiore a y (nel senso che $x \succeq_i y$ da tutti gli individui i e $x \succ_j y$ da qualche–almeno un–individuo j).

▲ Se rimuoviamo *iai*–come nel caso della regola di Borda–superiamo il problema della dittatorialità ad un costo aggiuntivo: l'ordinamento sociale di ogni coppia di alternative ora dipende anche da come gli individui ordinano tutte le altre alternative. Ciò inoltre comporta la possibilità di manipolare l'esito della votazione introducendo altre mozioni prima inesistenti.

▲ L'assioma di *transitività* della preferenza sociale debole ($\succeq_S$) ha come obiettivo quello di evitare ordinamenti ciclici (e il paradosso del voto). Tuttavia, rispetto a tale fine, essa è solo un'ipotesi sufficiente. Sen la sostituisce, prima, con l'ipotesi di transitività sulla preferenza sociale stretta ($\succ_S$), ma, come mostrato nel teorema di Gibbard, l'esito non è molto migliore di quello di Arrow in quanto la preferenza sociale individuata è *oligarchica*; dove, per oligarchica si intende che la preferenza sociale è definita dalle preferenze di un gruppo di individui di numero e identità incerti. In seguito, Sen sostituisce la transitività con l'assioma di *aciclicità* della preferenza sociale debole. Tuttavia, anche modificando alcuni altri assiomi (in particolare *iai*), egli perviene comunque a un nuovo risultato di "impossibilità".