Sistemi di scelte sociali:
la matematica è un'opinione

Introduzione

Come sappiamo, non esiste un meccanismo elettorale perfettamente equo dal punto di vista matematico. Se poi, invece di votare per un solo partito, i votanti devono stilare una graduatoria fra n partiti o candidati, è facile immaginare che, a maggior ragione, sarà facile incorrere in paradossi. Il fatto che le cose in effetti siano complicate è ben illustrato dal seguente esempio.

Esempio 1 Supponiamo che tre amici X, Y, Z debbano accordarsi per passare una serata, avendo a disposizione tre alternative: (a) andare al cinema, (b) andare a ballare, (c) giocare a carte. Se le rispettive preferenze, nell'ordine, sono quelle illustrate nella tabella seguente:

	X	Y	Z
I	a	b	c
II	b	c	a
III	c	a	b

allora i tre amici non possono mettersi d'accordo a maggioranza, perché due amici preferiscono a a b, due preferiscono b a c e due, infine, preferiscono c ad a. In pratica, le preferenze della maggioranza non rispettano la proprietà transitiva e quindi non è possibile ordinarle: questa circostanza è nota come il "paradosso del voto".
Ma non è finita qui: se i tre amici decidono di scegliere dapprima fra a e c, e successivamente fare un ballottaggio fra b e la vincente tra le due alternative precedenti, allora in prima istanza vince c coi voti di Y e Z, e nel ballottaggio vince b coi voti di X e Y. Ma se, viceversa, gli amici votano dapprima fra b e c, e successivamente fra a e la vincente, allora in primis vince b, votata da X e Y, e alla fine vince a, sostenuta da X e Z. E ancora, votando dapprima fra a e b vincerebbe a con il voto di X e Z, e poi fra a e c finirebbe col vincere c grazie a Y e Z. Quindi la scelta finale dipende dall'ordine con cui si votano le alternative. Non solo: se, ad esempio, nella votazione iniziale Z votasse per la sua seconda scelta a anziché per la sua prima scelta c, allora al primo

turno vincerebbe a col supporto di X e Z, mentre nel ballottaggio fra a e b vincerebbe a sostenuto ancora da X e Z; quindi passerebbe la seconda scelta di Z, invece della terza. In definitiva, può essere conveniente votare in modo "non sincero".

Ad una situazione del genere è legato anche il "paradosso dell'emendamento": consideriamo un parlamento con tre gruppi politici D, C e S, dove C è minoritario rispetto agli altri due, ma determinante (nel senso che sia D che S hanno la maggioranza assoluta solo con l'appoggio di C). Supponiamo che una certa legge L sia voluta da D e S, ma sia osteggiata da C. Se si dovesse scegliere fra l'approvazione di L e la sua bocciatura (alternativa NL), L sarebbe approvata poiché C è minoritario. Allora C propone un emendamento, che modifica L in LE, che è gradito a D ma fortemente sgradito a S, tanto da far sì che S preferisca l'alternativa NL rispetto a LE. La tabella delle alternative è allora simile alla precedente:

	D	C	S
I	LE	NL	L
II	L	LE	NL
III	NL	L	LE

Secondo le regole parlamentari, si vota dapprima l'emendamento, ossia fra L e LE: vince LE con i voti di D e C. A questo punto, il voto finale sulla legge, cioè fra LE e NL, vede prevalere NL coi voti di C e S. Così C, pur minoritario, è riuscito nel suo intento di bloccare la legge.

Relazioni di preferenza e di indifferenza

Dato un insieme K di alternative, introduciamo due relazioni binarie su K che serviranno a esprimere le preferenze di un individuo rispetto alle alternative disponibili: scrivendo $x\,P\,y$ intendiamo che *x è preferito* a y, mentre scrivendo $x\,I\,y$ intendiamo che le due alternative x e y sono *indifferenti*. Inoltre, le scritture $x\,\neg P\,y$ e $x\,\neg I\,y$ indicheranno rispettivamente che *x non è preferito* a y e che *x ed y non sono indifferenti*. In coerenza con il significato che attribuiamo alle relazioni P e I, stabiliamo per ipotesi le seguenti proprietà:

(1) se $x\,P\,y$ allora $y\,\neg P\,x$;

(2) se $x\,P\,y$ allora $x\,\neg I\,y$;

(3) se $x\,I\,y$ allora $x\,\neg P\,y$ e $y\,\neg P\,x$;

(4) vale una e una sola fra le relazioni $x\,P\,y$, $y\,P\,x$, $x\,I\,y$;

(5) se $x\,P\,y$ e $y\,P\,z$, allora $x\,P\,z$;

(6) se $x\,P\,y$ e $y\,I\,z$, allora $x\,P\,z$;

(7) se $x\,P\,y$ e $x\,I\,z$, allora $z\,P\,y$;

(8) se $x\,I\,y$ e $y\,I\,z$, allora $x\,I\,z$.

Le proprietà precedenti si riferiscono ad un soggetto "ideale": non sempre, nella realtà, le persone hanno preferenze transitive (5-6-7-8) o connesse (4).

Osservazione 2 Notiamo che la proprietà (1) implica che $x\,\neg P\,x$ per ogni x: quindi da (4) segue che $x\,I\,x$ per ogni x, ossia la relazione I è riflessiva. Inoltre, se $x\,I\,y$ si ha anche $y\,I\,x$: infatti, altrimenti per la (4) sarebbe $y\,P\,x$ oppure $x\,P\,y$; nel primo caso, essendo $y\,P\,x$ e $x\,I\,y$, la (6) implicherebbe $y\,P\,y$, assurdo per quanto osservato prima, mentre nel secondo caso, da $x\,P\,y$ e $x\,I\,y$ avremmo per la (7) $y\,P\,y$, nuovamente assurdo. Quindi la relazione I è simmetrica, nonché transitiva per la (8): dunque è una relazione di equivalenza. Invece la relazione P è irriflessiva, asimmetrica e transitiva, ossia è una relazione d'ordine stretto.
Una *graduatoria* è una sequenza di tutte le alternative di K, ordinate secondo la preferenza: alcune alternative possono essere tra loro indifferenti, ma la graduatoria ordina comunque tutte le alternative, raggruppandone le "classi di indifferenza".

Leggi di benessere sociale

Il problema che ci poniamo è il seguente: è dato un insieme C di n individui (i cittadini); ognuno di essi, diciamo l'h-simo, ha le sue personali relazioni di preferenza P_h e indifferenza I_h, con le quali formula la sua graduatoria personale. Ci chiediamo se sia sempre possibile, partendo da un qualsiasi insieme di n graduatorie individuali, definire due relazioni di preferenza P e di indifferenza I e formulare di conseguenza una graduatoria collettiva "accettabile". Ogni legge che associ ad una qualunque n-pla di graduatorie individuali una graduatoria collettiva è detta *legge di benessere sociale*.
L'esempio 1 mostra che la costruzione di una legge di benessere sociale soddisfacente non è un'impresa banale. Elenchiamo sotto forma di assiomi le

proprietà che, per semplice buonsenso, sembrano necessarie affinché una legge di benessere sociale sia realmente "democratica".

A1. Proprietà di completezza: *per ogni fissata* n*-pla di graduatorie individuali, la legge di benessere sociale determina una ed una sola graduatoria collettiva.*

Questo assioma dice semplicemente che la legge di benessere sociale è totale, ossia esprime una scala collettiva di preferenze, qualunque sia la varietà di preferenze espresse dai singoli. Dunque, i cittadini hanno piena libertà di scelta.

A2. Proprietà di sovranità: *per ogni coppia* x, y *di alternative distinte di* K *esiste una* n*-pla di graduatorie individuali tale che nella corrispondente graduatoria collettiva risulta* $x\,P\,y$*.*

Detto altrimenti, la legge di benessere sociale non ha preconcetti: per ogni coppia di alternative x, y può capitare che nella graduatoria collettiva sia $x\,P\,y$. In caso contrario non avremmo mai $x\,P\,y$, anche se tutti i cittadini preferissero x a y, e ciò violerebbe la sovranità dei cittadini.

A3: Proprietà di correlazione positiva: *se, per una coppia di alternative* $x, y \in K$ *e per una certa* n*-pla di graduatorie individuali, la legge di benessere sociale fornisce una graduatoria collettiva in cui* $x\,P\,y$*, allora per ogni altra* n*-pla di graduatorie individuali che differisce dalla precedente solo perché in alcune di esse* x *ha migliorato il suo piazzamento, nella graduatoria collettiva risultante deve valere* $x\,P\,y$*.*

Questo assioma assicura che vi è la necessaria correlazione positiva fra le graduatorie individuali e quella collettiva corrispondente: se nella graduatoria collettiva associata a una n-pla di graduatorie individuali risulta $x\,P\,y$, e se il gradimento di x fra i cittadini aumenta, a maggior ragione nella nuova graduatoria collettiva si avrà $x\,P\,y$.

A4. Proprietà di invarianza della alternative irrilevanti: *se, per una coppia di alternative* x, y *e per una certa* n*-pla di graduatorie individuali, la legge di benessere sociale fornisce una graduatoria collettiva in cui* $x\,P\,y$*, allora per ogni altra* n*-pla di graduatorie individuali in cui la relazione fra* x *e* y *non è mutata, nella graduatoria collettiva risultante deve ancora valere* $x\,P\,y$*.*

In altre parole, se la società collettivamente preferisce x a y, e se, pur modificandosi le graduatorie individuali, in tutte queste graduatorie x è ancora preferito a y, allora nella nuova graduatoria collettiva deve ancora aversi $x \, P \, y$. Quindi la legge di benessere sociale, per stabilire la preferenza fra due alternative x e y, non deve tener conto delle relazioni che x e y, nelle graduatorie individuali, hanno con tutte le altre alternative.

A5. Proprietà di non dittatorialità: *Non esiste alcun particolare cittadino $h \in C$, tale che, per ogni n-pla di graduatorie individuali, per ogni $x, y \in K$ risulti $x \, P \, y$ se e solo se $x \, P_h \, y$.*

Se tale individuo esistesse, egli sarebbe un dittatore, perché la graduatoria collettiva coinciderebbe sempre con la sua graduatoria personale.

Appare del tutto sorprendente che non esista alcuna legge di benessere sociale che verifichi questi cinque assiomi. Questo è ciò che esprime il teorema di Arrow, che andiamo ad esporre. Occorre sottolineare che Kenneth Arrow ha ottenuto per questo risultato il Premio Nobel per l'economia nel 1972, e che un altro premio Nobel per l'economia (1970), Paul Samuelson, ha paragonato l'impatto del teorema di Arrow sull'economia e sulla scienza politica a quello che ebbe il teorema di incompletezza di Gödel sulla logica e sui fondamenti della matematica.

Teorema (di Arrow) *Non esiste alcuna legge di benessere sociale che verifichi gli assiomi* **A1, A2, A3, A4, A5**.

Dimostrazione Cominciamo col mostrare che dagli assiomi **A2, A3** e **A4** discende la seguente *proprietà di Pareto o di unanimità*:

Lemma 1 (proprietà di unanimità) *Se, per una coppia di alternative x, y e per una n-pla di graduatorie individuali si ha $x \, P_h \, y$ per ogni $h \in \{1, \ldots, n\}$, allora per la corrispondente graduatoria collettiva deve essere $x \, P \, y$.*

Dimostrazione Sia data una n-pla di graduatorie individuali per cui si abbia $x \, P_h \, y$ per ogni $h \in \{1, \ldots, n\}$. Per **A4** possiamo limitarci a considerare le sole due alternative x e y, trascurando le altre; per **A2**, esiste un'altra n-pla di graduatorie individuali, le cui relazioni di preferenza indichiamo con P'_h, tale che nella corrispondente graduatoria collettiva si ha $x \, P' \, y$. Costruiamo adesso una nuova n-pla di graduatorie individuali, le cui relazioni di preferenza indichiamo con P''_h, modificando le preferenze P'_h in questo modo: le lasciamo inalterate se $x \, P'_h y$, e spostiamo x al primo posto in quelle in cui si ha $x \neg P'_h y$. Il risultato è che nella famiglia di preferenze individuali P''_h l'alternativa x ha migliorato la sua posizione rispetto a y; quindi, in virtù di **A3**, per la corrispondente preferenza collettiva P'' si deve avere $x \, P'' \, y$. Ma adesso, nelle preferenze individuali P''_h la situazione fra x e y è la stessa che si ha nelle preferenze individuali iniziali P_h (vale a dire x è preferito a y da tutti i cittadini). Dunque, tenuto conto che $x\,P''\,y$, per **A4** deve essere anche $x\,P\,y$. Ciò prova la tesi del Lemma 1.

Proseguiamo la dimostrazione del teorema di Arrow con una definizione.

Definizione 2 *Un insieme $D \subseteq C$ di individui si dice* decisivo per l'alternativa x rispetto all'alternativa y *se, per ogni n-pla di graduatorie individuali per le quali risulta $x\,P_h\,y$ per ogni $h \in D$, nella corrispondente graduatoria collettiva si ha $x\,P\,y$.*

Dunque D è decisivo se basta che tutti gli individui di D preferiscano x a y per essere sicuri che x venga preferito a y nella graduatoria collettiva.
Osserviamo che se un insieme decisivo D ha un solo elemento h, tale individuo è un dittatore "parziale", ossia detta legge solo per ciò che riguarda l'alternativa x rispetto all'alternativa y: se egli decide che $x\,P_h\,y$, allora certamente $x\,P\,y$.

Osservazione *Per ogni coppia di alternative $x, y \in K$ esiste almeno un insieme decisivo per x rispetto a y.*
Infatti, per la proprietà di unanimità, l'insieme C di tutti gli individui è decisivo per qualunque coppia di alternative.

Dimostriamo ora il seguente

Lemma 3 *Siano $x, y \in K$. Un sottoinsieme $D \subseteq C$ è decisivo per l'alternativa x rispetto all'alternativa y se e solo se risulta $x\,P\,y$ per ogni n-pla di graduatorie individuali per le quali si abbia $x\,P_h\,y$ per ogni $h \in D$, nonché $x\,\neg P_h\,y$ per ogni $h \in D^c$.*

Dimostrazione ($\Longrightarrow$) Sia D decisivo per l'alternativa x rispetto all'alternativa y. Allora, per definizione, risulta $x\,P\,y$ per tutte le n-ple di graduatorie individuali per le quali risulta $x\,P_h\,y$ per ogni $h \in D$; a maggior ragione ciò sarà vero per le n-ple che verificano sia $x\,P_h\,y$ per ogni $h \in D$, sia $x\,\neg P_h\,y$ per ogni $h \in D^c$. Dunque vale la condizione del lemma.
($\Longleftarrow$) Supponiamo che risulti $x\,P\,y$ per ogni n-pla di graduatorie individuali per le quali si abbia $x\,P_h\,y$ per ogni $h \in D$, nonché $x\,\neg P_h\,y$ per ogni $h \in D^c$. Per provare che D è decisivo per l'alternativa x rispetto all'alternativa y, consideriamo una n-pla di graduatorie individuali, le cui preferenze individuali denotiamo con P'_h, che verifichi $x\,P'_h\,y$ per ogni $h \in D$; occorre far vedere che

allora $x\,P'\,y$ per la graduatoria collettiva associata a tale n-pla. Consideriamo una nuova n-pla di graduatorie individuali, le cui preferenze indichiamo con P''_h, così fatta: per ogni cittadino $h \in D$ si definisce $P''_h = P'_h$; per ogni cittadino $h \notin D$, invece, spostiamo nell'ordine di preferenza x all'ultimo posto, cosicché si avrà $x \neg P''_h\, y$. In questo modo, nella nuova n-pla di graduatorie si ha $x\,P''_h y$ se e solo se $h \in D$ e $x \neg P''_h y$ se e solo se $h \notin D$. Dunque questa n-pla verifica la condizione del lemma, e pertanto $x\,P''\,y$ secondo la corrispondente preferenza collettiva. D'altra parte, la n-pla che esprime le preferenze P'_h si può pensare ottenuta da quella che esprime le preferenze P''_h migliorando i piazzamenti di x; tenuto conto che $x\,P''\,y$, per l'assioma **A3** deve risultare $x\,P'y$. Ciò prova che D è decisivo per x rispetto a y.

Proviamo adesso il lemma fondamentale, da cui il teorema di Arrow seguirà facilmente.

Lemma 4 *Se un sottoinsieme $D \subseteq C$ è decisivo per l'alternativa x rispetto all'alternativa y, allora D è decisivo per qualunque altra coppia ordinata di alternative.*

Notiamo che, di conseguenza, i dittatori "parziali" sono dittatori.

Dimostrazione Procediamo per casi.

(a) Fissiamo una coppia ordinata della forma (x, z), con $z \in K \setminus \{x, y\}$. Consideriamo una n-pla di alternative individuali tale che $x\,P_h\,z$ per ogni $h \in D$ e $x \neg P_h\,z$ per ogni $h \notin D$: vogliamo mostrare che $x\,P\,z$ secondo la preferenza collettiva associata a tale n-pla, e allora la decisività di D rispetto a (x, z) seguirà dal lemma 3. Costruiamo una nuova n-pla di graduatorie, le cui preferenze individuali indichiamo con P'_h, modificando la posizione di y (si noti che, per **A4**, ciò non altera le mutue posizioni che x e z hanno nella preferenza collettiva P): spostiamo y in modo che si abbia $x\,P'_h\,y$ e $y\,P'_h\,z$ per ogni $h \in D$, e in modo che si abbia $y\,P'_h\,z$ per $h \notin D$. Per la decisività di D rispetto a (x, y) deduciamo che $x\,P'\,y$, mentre per la proprietà di unanimità si ha $y\,P'\,z$, visto che vale $y\,P'_h\,z$ in tutto C. Dunque, per transitività, otteniamo $x\,P'\,z$ e, come si è detto, grazie all'assioma **A4** questa conclusione deve valere anche per la preferenza collettiva P: pertanto $x\,P\,z$, come si voleva.

(b) Fissiamo una coppia ordinata della forma (z, z'), con $z, z' \in K \setminus \{x\}$. Consideriamo una n-pla di alternative individuali tale che $z\,P_h\,z'$ per ogni $h \in D$ e $z \neg P_h\,z'$ per ogni $h \notin D$: vogliamo mostrare che $z\,P\,z'$, e la decisività di D rispetto a (z, z') seguirà dal lemma 3. Costruiamo una nuova n-pla di graduatorie, le cui preferenze individuali indichiamo con P'_h, modificando la posizione di x nel modo seguente: $z\,P'_h\,x$ e $x\,P'_h\,z'$ per ogni $h \in D$, $z\,P'_h\,x$ per ogni $h \notin D$. Poiché, per quanto dimostrato in (a), l'insieme D è decisivo per (x, z') (se $z' = y$ la decisività di D ce l'abbiamo per ipotesi), abbiamo $x\,P'\,z'$; per la proprietà di unanimità si ha poi $z\,P'\,x$, e quindi per transitività $z\,P'\,z'$. Dato che le mutue posizioni di z e z' non cambiano nel passaggio da P'_h a P_h, da **A4** segue allora che $z\,P\,z'$, come richiesto.

(c) Fissiamo una coppia ordinata della forma (z, x), con $z \in K \setminus \{x\}$. Consideriamo una n-pla di alternative individuali tale che $z\,P_h\,x$ per ogni $h \in D$ e $z\,\neg P_h\,x$ per ogni $h \notin D$: vogliamo mostrare che $z\,P\,x$, e come al solito la decisività di D rispetto a (z, x) seguirà dal lemma 3. Fissiamo un'alternativa $z' \in K \setminus \{x, z\}$ e costruiamo una nuova n-pla di graduatorie, le cui preferenze individuali indichiamo con P'_h, tale che risulti $z\,P'_h\,z'$ e $z'\,P'_h\,x$ per ogni $h \in D$, e $z'\,P'_h\,x$ per ogni $h \notin D$. Per quanto provato in (b), D è decisivo per (z, z') e quindi $z\,P'\,z'$; inoltre, per unanimità si ha $z'\,P'\,x$, cosicché per transitività ricaviamo $z\,P'\,x$. Dato che le mutue posizioni di x e z non cambiano nel passaggio da P'_h a P_h, da **A4** si ha $z\,P\,x$, come desiderato. Ciò conclude la dimostrazione del Lemma 4.

Dimostriamo finalmente il teorema di Arrow. Consideriamo una qualunque coppia di alternative (x, y). Come sappiamo, l'insieme C è certamente decisivo per (x, y); quindi l'insieme di tutti i sottoinsiemi di C che sono decisivi per (x, y) non è vuoto. Proviamo che esiste un sottoinsieme $D \subseteq C$ decisivo per (x, y) e *minimale*, ossia privo di sottoinsiemi propri non vuoti decisivi per (x, y). O è minimale l'insieme C stesso, oppure no: in questo caso, esisterà un suo sottoinsieme proprio non vuoto C', che è decisivo per (x, y). A sua volta, o C' è minimale, oppure esiste un sottoinsieme $C'' \subset C'$, non vuoto e decisivo per (x, y). Poiché C è un insieme finito, iterando all'estremo questo ragionamento si arriva ad un insieme D, decisivo per (x, y), costituito da un solo elemento $\{h\}$. Esso non ha sottoinsiemi propri non vuoti, quindi deve essere minimale: d'altra parte il lemma 4 ci dice che $D = \{h\}$ è decisivo per ogni coppia di alternative, ossia l'individuo h è un dittatore. Ciò però è assurdo, poiché contraddice l'assioma **A5**; ne segue che D deve avere almeno 2 elementi. Possiamo allora scrivere $D = D' \cup D''$, con D' e D'' disgiunti, non vuoti e, per minimalità, *non* decisivi per alcuna coppia di alternative. Fissiamo tre alternative x, y, z e consideriamo una n-pla di graduatorie individuali, definita dalle relazioni P_h, tale che risulti

$$\begin{aligned} &x\, P_h\, y \text{ e } y\, P_h z && \forall h \in D', \\ &z\, P_h\, x \text{ e } x\, P_h y && \forall h \in D'', \\ &y\, P_h\, z \text{ e } z\, P_h x && \forall h \notin D. \end{aligned}$$

Allora deve essere $x\, P\, y$, poiché D è decisivo per (x, y) e vale $x\, P_h\, y$ per ogni $h \in D$.
Consideriamo ora la coppia (y, z) e dimostriamo che $z \neg P\, y$. Siccome D'' non è decisivo per la coppia (z, y), esiste una n-pla di alternative, le cui preferenze individuali indichiamo con P'_h, tale che

$$z\, P'_h y \quad \forall h \in D'', \qquad z \neg P'_h\, y \quad \forall h \notin D'', \qquad z \neg P'\, y.$$

D'altra parte, riguardo alle preferenze P_h risulta per transitività $z\, P_h\, y$ per ogni $h \in D''$, mentre $z \neg P_h\, y$ per ogni $h \in D' \cup D^c = (D'')^c$. Quindi le mutue posizioni di z e y non cambiano fra P_h e P'_h, e di conseguenza, essendo $z \neg P'\, y$, **A4** ci dice che $z \neg P\, y$.
Proviamo adesso che $y \neg P\, z$. Supponiamo che sia $y\, P\, z$: intanto, dalla relazione già dimostrata $x\, P\, y$ seguirebbe $x\, P\, z$. Poi, dato che D' non è decisivo per la coppia (x, z), esiste una n-pla di alternative, le cui preferenze individuali indichiamo con P'_h, tale che

$$x\, P'_h z \quad \forall h \in D', \qquad x \neg P'_h\, z \quad \forall h \notin D', \qquad x \neg P'\, z.$$

Tuttavia, rispetto a P_h si hanno le stesse relazioni tra x e z: infatti, per transitività $x\, P_h z$ per ogni $h \in D'$, mentre $x \neg P_h\, z$ per ogni $h \in D'' \cup D^c = (D')^c$; dunque, come prima, da **A4** si ricava $x \neg P\, z$, il che contraddice la relazione $x\, P\, z$ dedotta sopra.
Proviamo infine che $y \neg I\, z$. Se fosse $y\, I\, z$, dalla relazione $x\, P\, y$ già nota segue subito $x\, P\, z$ e da qui, come prima, l'assurdo.
Abbiamo così provato che le alternative y e z non possono essere in alcun modo ordinate nella graduatoria collettiva, violando in tal modo l'assioma **A1**. Ciò è assurdo, il che significa che nessuna legge di benessere sociale può soddisfare simultaneamente gli assiomi **A1**, **A2**, **A3**, **A4** e **A5**.