L'assioma della scelta

Pino Vigna Suria

10 10 2004

1 Introduzione

L'assioma della scelta è una di quelle cose a cui tutti credono e che tutti usano senza accorgersene: non ho mai trovato nessuno studente che abbia fatto obiezioni alla semplice dimostrazione del fatto che ogni funzione suriettiva abbia un'inversa a destra; tale inversa viene trovata scegliendo un elemento in ognuno degli insiemi non vuoti costituiti dalle controimmagini degli elementi del codominio. Neanche i matematici fondatori della teoria degli insiemi avevano niente da ridire su questa procedura finché uno di loro, Zermelo, ha dimostrato nel 1904 che su ogni insieme si può mettere un ordine tale che ogni sottoinsieme non vuoto abbia un primo elemento. Questa è sembrata un po' grossa, tanto più che proprio pochi mesi prima che Zermelo tirasse fuori questa trovata König aveva dimostrato che su $\mathbb{R}$ un ordine con tale proprietà non esiste. Il vespaio che ne venne fuori convinse i matematici a guardare con più cura la dimostrazione di Zermelo, e uno di loro, Erhard Schmidt, si accorse che tale prova usava il principio di poter pescare un elemento in ogni insieme non vuoto anche se la famiglia di tali insiemi è infinita. Quando si rese conto di aver usato questo fatto (a cui aveva abboccato esattamente come gli studenti del primo anno abboccano alla dimostrazione dell'esistenza dell'inversa destra delle funzioni suriettive) Zermelo produsse un'altra dimostrazione del suo teorema citato sopra molto più semplice dell'originale. È la dimostrazione riportata in queste note. Per la cronaca König, stimolato dal trambusto, si accorse dell'errore della sua dimostrazione. Dopo questa erudita dissertazione storica veniamo alla matematica.

2 Prima serie di enunciati

Teorema 1. *I seguenti tre enunciati sono equivalenti:*

AC1- *Siano X un insieme non vuoto e $\mathcal{P}^*(X) = \mathcal{P}(X) - \{\emptyset\}$ l'insieme delle parti non vuote di X. Esiste una funzione, detta* funzione di scelta, *$f : \mathcal{P}^*(X) \longrightarrow X$ tale che $f(A) \in A,\ \forall A \in \mathcal{P}^*(X)$. In parole si può scegliere un elemento in ogni sottoinsieme non vuoto di X.*

AC2- *Ogni funzione suriettiva ha un'inversa a destra; cioè se A, B sono insiemi e $g : A \longrightarrow B$ è suriettiva esiste $h : B \longrightarrow A$ tale che $g \circ h(b) = b,\ \forall b \in B$.*

AC3- *Siano I, X insiemi non vuoti e $A : I \longrightarrow \mathcal{P}^*(X)$ una funzione; allora il prodotto cartesiano $\prod_{i\in I} A_i$ non è vuoto.*

Dimostrazione. Per dimostrare l'equivalenza seguiamo questo schema:

$$AC1 \Rightarrow AC2 \Rightarrow AC3 \Rightarrow AC1.$$

$AC1 \Rightarrow AC2$ è facile perché $g^{-1}(b) \in \mathcal{P}^*(X),\ \forall b \in B$ e, se f è una funzione di scelta, basta porre $h(b) = f(g^{-1}(b)),\ \forall b \in B$ e siamo a posto.

$AC2 \Rightarrow AC3$ Ricordiamo innanzitutto che

$$\prod_{i\in I} A_i = \{f : I \longrightarrow \bigcup_{i\in I} A_i \text{ tali che } f(i) \in A_i,\ \forall i \in I\}.$$

Definiamo $A'_i = A_i \times \{i\}$ e sia $g : \bigcup_{i\in I} A'_i \longrightarrow I$ definita da $g(x, i) = i$; g è suriettiva perché $\forall i \in I,\ A_i \neq \emptyset$ e quindi, per $AC2$, g ha un'inversa a destra h la quale soddisfa quindi $h(i) \in A'_i,\ \forall i \in I$; basta allora porre $f(i) = pr(g(i))$ dove $pr : \bigcup_{i\in I} A'_i \longrightarrow \bigcup_{i\in I} A_i$ è data da $pr(x, i) = x$ e siamo a posto.

$AC3 \Rightarrow AC1$ Grazie a $AC3$ sappiamo che $\prod_{A\in\mathcal{P}^*(X)} A \neq \emptyset$ cioè esiste $f : \mathcal{P}^*(X) \longrightarrow \bigcup_{A\in\mathcal{P}^*(X)} A = X$ tale che $f(A) \in A,\ \forall A \in \mathcal{P}^*(X)$. Ecco trovata una funzione di scelta. □

Chiameremo *assioma della scelta* e lo indicheremo con AC uno qualunque dei tre enunciati equivalenti appena visti.

Parliamo ora di insiemi ordinati. Sia $(X, \leq)$ un insieme parzialmente ordinato. Un sottoinsieme C non vuoto di X è detto *catena* se è totalmente ordinato dalla relazione $\leq$. Un elemento $x \in X$ viene detto *elemento massimale* se non esiste $y \in X$ tale che $y > x$. Se C è un sottoinsieme non vuoto di X e $x \in X$ si dice che x è un *maggiorante* di C se, per ogni $y \in C$, $y \leq x$; si dice che $a \in C$ è il *primo elemento* di C se, per ogni $y \in C$, $a \leq y$. Per esempio se A è un insieme non vuoto e $X = \mathcal{P}(A) - \{A\}$ è ordinato mediante l' inclusione allora gli insiemi della forma $A - \{a\}$ sono tutti e soli gli elementi massimali. L'insieme vuoto è il primo elemento di X. Il primo elemento di $\mathbb{N}$ con il solito ordine è 0 mentre $\mathbb{Z}$ e l'insieme dei numeri reali positivi non hanno primo elemento. 21 è un maggiorante per il sottoinsieme di $\mathbb{Z}$ formato dai numeri interi negativi. Ogni sottoinsieme di un insieme totalmente ordinato è una catena. $(X, \leq)$ è detto *induttivo* se ogni sua catena ha un maggiorante. $(X, \leq)$ è detto *bene ordinato* e $\leq$ un *buon ordine* su X se ogni sottoinsieme non vuoto di X ha un primo elemento.

Teorema 2. *I seguenti enunciati sono equivalenti tra di loro ed all'assioma della scelta:*

- *BO* *Dato un insieme non vuoto X esiste un buon ordine su X (questo è il teorema di Zermelo che ha fatto nascere il caso).*
- *LZ* *(Lemma di Zorn)- Ogni insieme induttivo ha un elemento massimale.*
- *TH* *(Teorema di Hausdorff)- Sia $(X, \leq)$ un insieme parzialmente ordinato e $\mathcal{C}$ l'insieme delle catene di X ordinato mediante l'inclusione; esiste un elemento massimale $C^* \in \mathcal{C}$.*

Dimostrazione. Schema della dimostrazione:

$$AC \Rightarrow BO \Rightarrow AC$$

$$BO \Rightarrow TH \Rightarrow LZ \Rightarrow BO$$

Cominciamo con le implicazioni più facili:

$BO \Rightarrow AC$. Sia X un insieme non vuoto e $\leq$ un buon ordinamento su di lui. La funzione $f : \mathcal{P}^*(X) \longrightarrow X$ data da $f(A) =$ primo elemento di A è una funzione di scelta.

$LZ \Rightarrow BO$. Sia X un insieme non vuoto e sia $\mathcal{X} = \{(A, \leq_A)\}$ dove $A \in \mathcal{P}^*(X)$ e $\leq_A$ è un buon ordine su A. $\mathcal{X}$ non è non vuoto perché i singoletti, muniti dell'unico ordine possibile, ci appartengono. Definiamo una relazione d'ordine $\preceq$ su $\mathcal{X}$ come segue $(A, \leq_A) \preceq (B, \leq_B)$ se e solo se $A \subseteq B$, $\leq_A$ è la restrizione ad A di $\leq_B$ e $\forall a \in A, b \in B - A,\ a \leq_B b$.

$(\mathcal{X}, \preceq)$ è induttivo; sia infatti $\mathcal{C}$ una catena di $\mathcal{X}$ e poniamo

$$A = \bigcup_{(C, \leq_C) \in \mathcal{C}} C.$$

Su A definiamo un ordine $\leq_A$ così: se $x, y \in A$, dato che $\mathcal{C}$ è una catena, esiste $C \in \mathcal{C}$ tale che $x, y \in C$; poniamo allora $x \leq_A y$ se e solo se $x \leq_C y$. Questa definizione non dipende dalla scelta di C infatti se B è un altro elemento di $\mathcal{C}$ tale che $x, y \in B$ allora $B \subseteq C$ (o viceversa) e $x \leq_B y$ se e solo se $x \leq_C y$. $\leq_A$ è un buon ordine su A infatti se $K \subseteq A$ non è vuoto esiste $C \in \mathcal{C}$ tale che $K \cap C$ non è vuoto; se c è il primo elemento di $K \cap C$ secondo l'ordine $\leq_C$ allora $\forall x \in K$, o $x \in C$ a allora $c \leq_C x$ e quindi $c \leq_A x$ oppure $x \notin C$, nel qual caso esiste $B \in \mathcal{C}$ tale che $x \in B$; non può essere $B \subseteq C$ e quindi deve essere $C \subseteq B$ e allora $c \leq_B x$ e quindi $c \leq_A x$; c è dunque il primo elemento di K secondo $\leq_A$. Perciò $(A, \leq_A) \in \mathcal{X}$ è un maggiorante di $\mathcal{C}$; per LZ $\mathcal{X}$ ha un elemento massimale $(A^*, \leq_{A^*})$; dico che deve essere $A^* = X$ e in questo modo $\leq_{A^*}$ sarà un buon ordine su X. Infatti se $a \in X - A^*$ definiamo $A' = A^* \cup \{a\}$ e $\leq_{A'}$ così: $x \leq_{A'} y$ se e solo se $x \leq_{A^*} y$ quando $x, y \in A^*$ e $x \leq_{A'} a$, $\forall x \in A^*$. Evidentemente $\leq_{A'}$ è un buon ordinamento su A' e quindi $(A', \leq_{A'}) \in \mathcal{X}$. Inoltre $(A^*, \leq_{A^*}) \prec (A', \leq_{A'})$ contro la massimalità di $(A^*, \leq_{A^*})$ in $\mathcal{X}$. $\square$

Prima di procedere alle altre implicazioni bisogna introdurre un po' di terminologia. Sia $(X, \leq)$ un insieme parzialmente ordinato e $a \in X$; il *segmento* di a è l'insieme $X_a = \{x \in X \text{ tali che } x < a\}$.

Teorema 3 (principio di induzione trasfinita)**.** *Sia $(X, \leq)$ un insieme bene ordinato. Se $U \subseteq X$ gode della seguente proprietà*

$$\forall a \in X,\ X_a \subseteq U \Rightarrow a \in U$$

allora $U = X$

Dimostrazione. Sia, per assurdo, $X - A \neq \emptyset$ e a il suo primo elemento; questo significa che $a \notin U$ ma che ogni $x < a$ appartiene a U, cioè precisamente che il segmento $X_a \subseteq U$ ma $a \notin U$, contraddicendo la proprietà. $\square$

Questa induzione trasfinita somiglia alla solita induzione su $\mathbb{N}$ che è un insieme bene ordinato dal solito ordine; la differenza essenziale è che in $\mathbb{N}$ ogni elemento non zero è il successore di un altro elemento, e questo non è vero in generale benché si possa ancora definire l'idea di successore di un elemento $a \in X$ come il primo elemento di $\{x \in X$ tali che $x > a\}$ (se questo non è vuoto). Il seguente esempio illustrerà la situazione: su $\mathbb{N} \times \mathbb{N}$ definiamo un buon ordine $\preceq$ ponendo $(a, b) \preceq (c, d)$ se e solo se $a \leq c$ o $a = c$ e $b \leq d$ (una specie di ordine alfabetico sulle parole composte da due numeri naturali) tutti gli elementi della forma $(0, a)$ non sono il successore di nessun elemento. Per analogia con il caso noto l'ipotesi $X_a \subseteq U$ verrà chiamata *ipotesi induttiva.*

AC $\Rightarrow$ BO.

Dimostrazione. Sia X un insieme non vuoto e sia $f : \mathcal{P}^*(X) \longrightarrow X$ una funzione di scelta. Un *f-insieme* è una coppia $(A, \leq_A)$ dove $A \in \mathcal{P}^*(X)$ e $\leq_A$ è un buon ordine su A tale che ,$\forall a \in A$ si abbia $a = f(X - A_a)$ (non é una richiesta assurda perché $a \in X - A_a$ che è quindi non vuoto). Per esempio sia $m = f(X)$; il singoletto $\{m\}$ con l'unico ordine che gli si può dare è un $f-$ insieme perchè $\{m\}_m = \emptyset$ e quindi $f(X - \{m\}_m) = f(X - \emptyset) = f(X) = m$.

Proponiamo qualche considerazione che, pur non essenziale per la dimostrazione, ci darà una certa confidenza con gli $f-$insiemi che sarà utile nella dimostrazione vera e propria. Sia $(A, \leq_A)$ un $f-$insieme e sia a_0 il suo primo elemento (secondo l'ordine $\leq_A$) allora $A_{a_0} = \emptyset$ e dunque $a_0 = f(X - A_{a_0}) = f(X - \emptyset) = f(X) = m$. Quindi m è forzatamente il primo elemento di ogni $f-$insieme, il che già ci dice che ci sono forti limitazioni su $\leq_A$. Ma non finisce qui: sia a_1 il primo elemento di $A - \{a_0\}$ (se questo non è vuoto); allora $a_1 = f(X - A_{a_1}) = f(X - \{a_0\}) = f(X - \{m\}) = m_1$. E così via; sembra che gli elementi di un $f-$insieme siano del tutto forzati insieme con il loro ordine.È proprio così se appena riusciamo a capire cosa vogliamo dire. Un segmento di un $f-$insieme è ancora un $f-$insieme perché se $b \in A_a$ allora $X - (A_a)_b = X - A_b$.

Se $(A, \leq_A)$ e $(B, \leq_B)$ sono due $f-$insiemi tali che $A \subseteq B$ allora $\forall a \in A, A_a = B_a$ (il che in particolare dice che $\leq_{B|A} = \leq_A$ e c'è un solo ordine possibile su un insieme che lo renda un $f-$insieme). Dimostriamo questo fatto per induzione trasfinita; sia $U = \{a \in A$ tali che $A_a = B_a\}$. Sia allora $A_a \subseteq U$ e proviamo che $a \in U$. Sia b il primo elemento di $B - A_a$ (ovviamente secondo l'ordine $\leq_B$) che non è vuoto perché a ci appartiene. Dico che

$A_a = B_b$; infatti se $x \in A_a$ allora $x \in U$ e quindi $A_x = B_x$, per cui $B_x \subseteq A_a$ e siccome $b \notin A_a$ avremo che $b \notin B_x$ e cioè $b \geq_B x$; inoltre $x \notin B - A_a$ e quindi $x \neq b$; ne segue che quindi $x \in B_b$. Viceversa sia $x \in B_b$; allora $x \notin B - A_a$ in quanto b è il primo elemento di tale insieme; perciò $x \in A_a$. Ma $a = f(X - A_a) = f(X - B_b) = b$.

Ho dimostrato che $a \in U$ e per il principio d'induzione trasfinita $U = A$, che è quanto si voleva provare. Se $A \subset B$ allora sia b il primo elemento di $B - A$. Dico che $A = B_b$. Infatti se $x \in B_b$ allora $x <_B b$ e quindi $x \notin B - A$ dato che b è il primo elemento di $B - A$. Viceversa se $x \notin B_b$ allora $x \geq_B b$ e quindi o $x = b$ il che ci garantisce che $x \notin A$ o $b \in B_x$; se fosse $x \in A$ allora $B_x = A_x$ e quindi $b \in A_x \subset A$; assurdo. Abbiamo fatto vedere che due $f-$insiemi uno contenuto nell'altro o coincidono (ordine compreso) o uno è un segmento dell'altro. Lo stesso è vero anche se non si suppone a priori che siano uno contenuto nell'altro. Siano infatti $(A, \leq_A)$ e $(B, \leq_B)$ due f-insiemi tali che A non sia contenuto in B; allora $A - (A \cap B) \neq \emptyset$ e sia a il suo primo elemento (secondo, ovviamente, $\leq_A$); $A_a \subseteq B$ perché se $x <_A a$ allora $x \notin A - (A \cap B)$ dato che a è il primo elemento di questo insieme e quindi $x \in B$. A_a è dunque un $f-$insieme contenuto nell'$f-$insieme B; quindi $A_a = B$ (e quindi B è un segmento di A) oppure esiste un $b \in B$ tale che $A_a = B_b$; in quest'ultimo caso si vede come sopra che $a = b$ il che è assurdo perché $a \notin B$.

Abbiamo provato che due $f-$insiemi o coincidono (ordine compreso) o uno è un segmento dell'altro.

Sia $\mathcal{F}$ l'insieme di tutti gli $f-$insiemi di X e poniamo $L = \bigcup_{(A, \leq_A) \in \mathcal{F}} A$. Poniamo il L un ordine $\leq_L$ così: se $x, y \in L$ allora esistono due $f-$insiemi $(A, \leq_A)$ e $(B, \leq_B)$ tale che $x \in A$ e $y \in B$; per quanto abbiamo visto possiamo supporre che $A \subseteq B$; decidiamo che $x \leq_L y$ se e solo se $x \leq_B y$. È facile vedere che questa definizione non dipende dalla scelta dell'$f-$insieme B in cui ci sono sia x che y (cfr. dimostrazione del fatto che $LZ \Rightarrow BO$).

$\leq_L$ è un buon ordine su L perché se K è un sottoinsieme non vuoto di L allora K interseca almeno un $f-$insieme A. Sia a il primo elemento di $A \cap K$ (secondo $\leq_A$); se b è un qualunque altro elemento di K, allora o $b \in A$ e perciò $a \leq_A b$ e quindi $a \leq_L b$ oppure $b \notin A$ e appartiene invece a un altro $f-$insieme $B \neq A$; è impossibile che B sia un segmento di A e quindi A è un segmento B_c di B; quindi $a \in B$. Ma allora $a <_B c \leq_B b$ dato che $b \notin B_c = A$. Quindi $a \leq_B b$ e $a \leq_L b$. Cioè a è il primo elemento di K (secondo $\leq_L$).

$(L, \leq_L)$ è un $f-$insieme. infatti se $a \in L$ allora $a \in A$, dove A è un $f-$insieme. Dico che $L_a = A_a$ dal che si deduce che $a = f(X - A_a) =$

$f(X - L_a)$; che A_a sia contenuto in L_a è evidente; viceversa se $b \in L_a$ allora o $b \in A$ e allora siamo a posto o $b \in B - A$ dove B è un altro $f-$insieme; visto che A non contiene B, A deve essere un segmento B_c di B; ma allora $b \notin B_c$ e dunque $c \leq_B b$ e $a <_B b$ e perciò $a \leq_L b$ e quindi $b \notin L_a$; assurdo.

Concludiamo facendo vedere che $L = X$, in tal caso $\leq_L$ sarà il buon ordine che cercavamo per X. Sia per assurdo $X - L \neq \emptyset$ a sia $t = f(X - L)$; poniamo $L' = L \cup \{t\}$ e definiamo su L' un ordine $\leq_{L'}$ decidendo che ristretto a L sia $\leq_L$ e $t \geq_{L'} x$, $\forall x \in L'$. Chiaramente $\leq_{L'}$ è un buon ordine e $(L', \leq_{L'})$ è ancora un $f-$insieme: infatti se $a \in L'$ o $a \in L$ e allora $a = f(X - L_a) = f(X - L'_a)$ o $a = t$ e allora $t = f(X - L) = f(X - L'_t)$. Abbiamo così trovato un $f-$insieme propriamente più grande di tutti gli $f-$insiemi. Assurdo. $\square$

Osserviamo che l'induzione trasfinita può essere usata, oltre che per dimostrare delle affermazioni, per definire dei sottinsiemi di un insieme bene ordinato $(X, \leq)$; infatti, come si impara al primo giorno di scuola, conoscere un sottoinsieme A di X vuol dire poter decidere per ogni elemento $x \in X$ se $x \in A$ o $x \notin A$. Poniamo allora $U = \{x \in X$ tali che sappiamo decidere se $x \in A$ o $x \notin A\}$; conosceremo A precisamente quando $U = X$ e per provare questo si può usare l'induzione trasfinita cioè decidere se $a \in A$ quando si sa già, per tutti gli $x < a$, se questi stanno o no in A.

Usando questa tecnica proviamo che

$BO \Rightarrow TH$

Sia $(X, \leq)$ un insieme parzialmente ordinato. Mettiamo su X un buon ordine $\preceq$ (attenzione: $\preceq$ non ha niente a che fare con $\leq$). Mediante l'induzione trasfinita separiamo X in due classi disgiunte V e W così: Mettiamo il primo elemento di X in V; supponiamo poi che per un certo $a \in X$ si sia deciso, per ogni $x \prec a$ se $x \in V$ o $x \in W$; in altri termini gli insiemi V_a e W_a formati rispettivamente dagli elementi di X_a che stanno in V o in W sono noti. Mettiamo a in V se a é confrontabile con tutti gli elementi di $A \cap V_a$ e altrimenti lo mettiamo in W. Dico che V è una catena massimale: è una catena perché se $x, y \in V$ e $x \prec y$ allora $x \in V \cap X_y$ e quindi x e y sono confrontabili; ed è massimale perché se V' è una catena che contiene V e $y \in V'$ allora, per ogni $x \in V \cap X_y \subseteq V' \cap X_y$, y è confrontabile con x, e quindi $y \in V$.

$TH \Rightarrow LZ$

Siano $(X, \leq)$ un insieme induttivo, C una sua catena massimale (che esiste per TH) e c un maggiorante di C (che esiste perché X è induttivo). Se esiste $x \in X$ tale che $x > c$ allora la catena $C \cup \{x\}$ è una catena che contiene propriamente C, contro la sua massimalità.