APPUNTI DI TEORIA DEGLI INSIEMI

MAURIZIO CORNALBA

L'assioma della scelta e il lemma di Zorn

Sia $\{A_i\}_{i\in I}$ un insieme di insiemi. Il prodotto $\prod_{i\in I} A_i$ è l'insieme di tutte le applicazioni $\alpha : I \to \cup_{i\in I} A_i$ tali che $\alpha(i) \in A_i$ per ogni i; a volte si scrive α_i per indicare $\alpha(i)$ e si indica la funzione α con $\{\alpha_i\}_{i\in I}$. L'*assioma della scelta*, in una delle sue molte forme, dice quanto segue:

Assioma della scelta. *Se $\{A_i\}_{i\in I}$ è un insieme non vuoto di insiemi non vuoti, il prodotto $\prod_{i\in I} A_i$ è non vuoto.*

Una risultato standard per dimostrare il quale è necessario usare l'assioma della scelta è il seguente.

Proposizione 1. *Una applicazione $f : A \to B$ è suriettiva se e solo se esiste una applicazione $g : B \to A$ tale che $f \circ g$ sia l'applicazione identica di B in sè.*

Che f sia suriettiva se esiste una g con la proprietà descritta nel lemma è ovvio (e non richiede l'uso dell'assioma della scelta). Viceversa, se poniamo $A_b = f^{-1}(b)$, dire che f è suriettiva equivale a dire che A_b non è vuoto per ogni $b \in B$. L'assioma della scelta dice allora che il prodotto degli A_b non è vuoto. Qualsiasi elemento g di questo prodotto ha le proprietà richieste.

Vi sono vari altri enunciati equivalenti all'assioma della scelta; ne dimostreremo alcuni. Ricordiamo che un *semiordinamento* su un insieme X è una relazione $\leq$ su X tale che:

(1) se $x \in X$, allora $x \leq x$;
(2) se $x \leq y$ e $y \leq x$, allora $x = y$;
(3) se $x \leq y$ e $y \leq z$, allora $x \leq z$.

Come è naturale, scriveremo $x \geq y$ per indicare che $y \leq x$, e $x < y$ (o $y > x$) per indicare che $x \leq y$ ma $x \neq y$. Un *insieme semiordinato* è un insieme con un semiordinamento. Sia X un insieme semiordinato. Diremo che due elementi $x, y \in X$ sono *confrontabili* se $x \leq y$ o $x \geq y$. Un insieme semiordinato è *totalmente ordinato* (o semplicemente *ordinato*) se due suoi qualsiasi elementi sono confrontabili. Diremo che $x \in X$ è un *massimo* se $x \geq y$ per ogni $y \in X$, che è un elemento *massimale* se non vi sono elementi $y \in X$ tali che $x < y$. Se Y è un sottoinsieme di X, un *maggiorante* di Y è un elemento x di X tale che $x \geq y$ per ogni $y \in Y$. In modo analogo si possono definire le nozioni di *minimo*, elemento *minimale* e *minorante*. Una *catena* in X è un sottoinsieme totalmente ordinato di X. Se C è una catena in X, l'*estremo superiore* di C (se esiste) è il minimo tra gli $x \in X$ tali che $x \geq c$ per ogni $c \in C$; in modo analogo si definisce la nozione di *estremo inferiore*. Diremo che un insieme semiordinato X è *induttivo* se ogni catena in X ha un maggiorante. Il risultato che segue va sotto il nome di "lemma di Zorn".

Teorema 1. (Lemma di Zorn) *Ogni insieme semiordinato, non vuoto e induttivo contiene elementi massimali.*

Dedurremo il lemma di Zorn da un altro risultato notevole, che va sotto il nome di "principio massimale di Hausdorff". Prima di enunciarlo notiamo che l'insieme di tutte le catene in un insieme semiordinato è a sua volta un insieme semiordinato (per inclusione).

Teorema 2. *Ogni insieme semiordinato contiene catene massimali.*

Deduciamo il lemma di Zorn da questo risultato. Se X è semiordinato, contiene una catena massimale C. Se X è induttivo, questa catena ha un maggiorante m. Se m non fosse un elemento massimale di X, esisterebbe $x \in X$ tale che $x > m$. Ma allora $C \cup \{x\}$ sarebbe una catena contenente strettamente C, contro la massimalità di quest'ultima.

Per dimostrare il principio massimale di Hausdorff ci appoggeremo su un lemma tecnico. Se x e y sono elementi di un insieme semiordinato diremo che y è un *successore* di x se $y > x$ e non ci sono elementi z tali che $y > z > x$.

Lemma 1. *Sia X un insieme semiordinato. Supponiamo che X abbia minimo e che ogni catena in X abbia estremo superiore. Allora esistono elementi di X che non hanno successori.*

Il principio massimale di Hausdorff segue facilmente dal lemma. Sia infatti X un insieme semiordinato e $\mathcal{Z}$ l'insieme delle catene in X. L'insieme $\mathcal{Z}$ è a sua volta semiordinato per inclusione, e possiede un minimo, cioè l'insieme vuoto. Sia $\mathcal{C}$ una catena in $\mathcal{Z}$. L'insieme $\cup\mathcal{C}$ è una catena in X. Infatti se x, y sono elementi di $\cup\mathcal{C}$, esistono elementi A e B di $\mathcal{C}$ tali che $x \in A$ e $y \in B$. Poiché $\mathcal{C}$ è una catena in Z, $A \subset B$ oppure $B \subset A$. Dunque x e y sono entrambi contenuti in A o in B. Dato che A e B sono catene, x e y sono confrontabili. Inoltre $\cup\mathcal{C}$ è l'estremo superiore di $\mathcal{C}$. Infatti una catena in X che contenga ogni elemento di $\mathcal{C}$ deve necessariamente contenere $\cup\mathcal{C}$. Il lemma 1 dice allora che esiste una catena D in X che non possiede successori. Una tale catena deve essere necessariamente massimale. Se infatti esistesse una catena E che la contiene strettamente, se vi fosse cioè un elemento $x \in E$ che non appartiene a D, la catena $D \cup \{x\}$ sarebbe un successore di D.

Resta da dimostrare il lemma 1. Ragioneremo per assurdo, supponendo che ogni elemento di X abbia un successore, supponendo cioè che per ogni $x \in X$ l'insieme dei successori di x sia non vuoto. L'assioma della scelta dice allora che esiste una funzione f che associa ad ogni $x \in X$ un suo successore $f(x)$. Indichiamo con p il minimo di X. Chiameremo p-successione un sottoinsieme Y di X con le seguenti proprietà:

(1) $p \in Y$;
(2) se $x \in Y$, allora $f(x) \in Y$;
(3) se C è una catena in Y, il suo estremo superiore appartiene a Y.

Osserviamo innanzitutto che esistono p-successioni, dato che X stesso è una p-successione. Notiamo poi che l'intersezione di una famiglia di p-successioni è una p-successione. Dunque l'intersezione di tutte le p-successioni, che indichiamo con P, è una p-successione. Chiameremo un elemento di P *privilegiato* se è confrontabile con ogni elemento di P. Per ogni x privilegiato poniamo

$$A_x = \{y \in P : y \leq x \text{ oppure } y \geq f(x)\}.$$

Dico che A_x è una p-successione, e quindi è uguale a P, dato che è in esso contenuta. In altre parole, se x è privilegiato, per ogni $y \in P$ si ha che $y \leq x$ o $y \geq f(x)$.

Innanzitutto $p \in A_x$. Se $y \in A_x$, per mostrare che $f(y) \in A_x$ distinguiamo vari casi. Se $y < x$, $f(y) \leq x$ dato che x è confrontabile con ogni elemento di P, e dunque in particolare con $f(y)$. Se $y = x$, allora $f(y) \geq f(x)$. Se $y \geq f(x)$, allora $f(y) > y \geq f(x)$. In ogni caso si conclude che $f(y) \in A_x$. Sia infine C una catena in A_x, e sia m il suo estremo superiore. Se esiste $y \in C$ tale che $y \geq f(x)$, allora $m \geq y \geq f(x)$, e quindi $m \in A_x$. Altrimenti $y \leq x$ per ogni $y \in C$, e dunque $m \leq x$. In ogni caso, $m \in A_x$. Ciò dimostra quanto affermato, cioè che A_x è una p-successione.

Una conseguenza di quanto abbiamo mostrato è che l'insieme di tutti gli elementi privilegiati di P è una p-successione. In effetti, se x è privilegiato e y è un elemento qualsiasi di P, sappiamo che o $y \leq x < f(x)$, oppure $y \geq f(x)$. Dunque $f(x)$ è confrontabile con ogni elemento di P, cioè è privilegiato. Se invece C è una catena di elementi privilegiati, m è il suo estremo superiore, e y è un elemento di P, possiamo distinguere due casi. O esiste $x \in C$ tale che $x \geq y$, nel qual caso $m \geq x \geq y$, oppure $x \leq y$ per ogni $x \in C$, e quindi $m \leq y$; in ogni caso si conclude che m è confrontabile con ogni elemento di P. Poiché l'insieme degli elementi privilegiati di P è una p-successione, deve coincidere con P; ne segue in particolare che P è totalmente ordinato.

A questo punto possiamo concludere la dimostrazione del lemma 1 raggiungendo una contraddizione. Poiché P è una catena in X, ha estremo superiore m per ipotesi; dato che P è una p-successione, m appartiene a P. Allora, sempre perchè P è una p-successione, $f(m) \in P$. Ma questo è assurdo, perchè $f(m) > m$ e m è l'estremo superiore di P.

Esercizio 1. Dimostrare che il lemma di Zorn è equivalente all'assioma della scelta. In altre parole, mostrare che se si assume come assioma il lemma di Zorn se ne può dedurre l'assioma della scelta nella forma enunciata all'inizio di queste note.

Esercizio 2. Dimostrare che il principio massimale di Hausdorff (Teorema 2) segue dal lemma di Zorn.

L'assioma della scelta ha conseguenze sorprendenti. Una di queste è il teorema di buon ordinamento. Un *buon ordinamento* su un insieme X è un ordinamento con la proprietà che ogni sottoinsieme non vuoto di X ha minimo. Un insieme *bene ordinato* è un insieme munito di un buon ordinamento. Ad esempio $\mathbb{N}$, con l'ordinamento usuale, è bene ordinato; questo equivale al principio di induzione.

Teorema 3. (del buon ordinamento) *Ogni insieme ammette un buon ordinamento.*

La dimostrazione che daremo usa il lemma di Zorn. Sia X un insieme non vuoto, e sia $\mathcal{Z}$ l'insieme delle coppie $a = (I_a, \leq_a)$, dove I_a è un sottoinsieme di X e $\leq_a$ è un buon ordinamento su I_a. Introduciamo un semiordinamento su $\mathcal{Z}$ ponendo $a \leq b$ se $I_a \subset I_b$, la restrizione di $\leq_b$ ad I_a è $\leq_a$, e inoltre $x \leq_b y$ ogni volta che $x \in I_a$, $y \in I_b$ ma $y \notin I_a$. L'insieme $\mathcal{Z}$ non è vuoto, perchè ogni insieme finito ammette un buon ordinamento. Dico che $\mathcal{Z}$ è induttivo. Sia $\mathcal{C}$ una catena in $\mathcal{Z}$. Poniamo $A = \cup_{a \in \mathcal{C}} I_a$, e consideriamo gli ordinamenti $\leq_a$ come relazioni su A, cioè come sottoinsiemi di $A \times A$. Poniamo poi $\leq_A = \cup_{a \in \mathcal{C}} \leq_a$. È chiaro che $\leq_A$ è un ordinamento totale su A che, per ogni $a \in \mathcal{C}$, induce l'ordinamento $\leq_a$ su I_a; mostriamo che è anche un buon ordinamento. Sia D un sottoinsieme non vuoto di A, e sia x un suo elemento. Allora esiste $a \in \mathcal{C}$ tale che $x \in I_a$. Se y è un elemento di A che precede x, esiste $b \in \mathcal{C}$ tale che $y \in I_b$. Se $b \leq a$, y appartiene a I_a. Se invece $a \leq b$, y deve comunque

appartenere ad I_a altrimenti, per la definizione dell'ordinamento su $\mathcal{Z}$, seguirebbe x. Ora $I_a \cap D$ è un sottoinsieme di I_a, e quindi ha minimo. Per quanto si è appena osservato, questo minimo è un minimo anche per D. È ora chiaro che $(A, \leq_A)$ è l'estremo superiore di $\mathcal{C}$. Dunque $\mathcal{Z}$ è induttivo, e quindi ammette un elemento massimale $(F, \leq_F)$. Se F fosse strettamente contenuto in X, cioè se vi fosse un elemento x di X non appartenente a F, potremmo estendere $\leq_F$ a un ordinamento di $F \cup \{x\}$ imponendo a x di seguire ogni elemento di F. Questo sarebbe un buon ordinamento su $F \cup \{x\}$, contro la massimalità di $(F, \leq_F)$. In conclusione, $\leq_F$ è un buon ordinamento su $X = F$.

Esercizio 3. Mostrare che il teorema di buon ordinamento è equivalente all'assioma della scelta.

Cardinalità

Si dice che due insiemi sono *equipotenti* se si possono mettere in corrispondenza biunivoca. Un insieme si dice *finito* se è vuoto o equipotente a un insieme della forma $\{1, \dots, n\}$ per qualche intero positivo n, *infinito* se non è finito. Un insieme si dice *numerabile* se è equipotente a $\mathbb{N}$.

Teorema 4. *Sia X un insieme. Le seguenti condizioni sono equivalenti:*
i) X è infinito;
ii) X contiene un sottoinsieme numerabile;
iii) X è equipotente a un suo sottoinsieme proprio.

Mostriamo che *i)⇒ii)*. Scegliamo un elemento x_1 di X. Dato che X è infinito, è diverso da $\{x_1\}$. Sia x_2 un elemento di X diverso da x_1. Dato che X è infinito, è diverso da $\{x_1, x_2\}$. Sia x_3 un elemento di X non appartenente a $\{x_1, x_2\}$. Procedendo in questo modo si costruisce una successione $\{x_i : i = 1, 2, \dots\}$ di elementi distinti di X, cioè un sottoinsieme numerabile di X.

Mostriamo che *ii)⇒iii)*. Sia Y un sottoinsieme numerabile di X, e sia $\alpha : \mathbb{N} \to Y$ una applicazione biunivoca. Definiamo una applicazione β da X in sè ponendo $\beta(x) = x$ se $x \notin Y$ e $\beta(\alpha(n)) = \alpha(n+1)$. Questa applicazione è iniettiva ma non suriettiva perchè $\alpha(1)$ non appartiene alla sua immagine. Dunque X è equipotente a $X \setminus \{\alpha(1)\}$.

Mostriamo infine che se non vale *i)* non vale nemmeno *iii)*. Mostriamo cioè che un insieme finito non è equipotente a una sua parte propria. Supponiamo che ciò sia falso, e sia n il minimo intero tale che $\{1, \dots, n\}$ sia equipotente a un suo sottoinsieme proprio. Stiamo dunque supponendo che esista una applicazione α di $\{1, \dots, n\}$ in sè che è iniettiva ma non suriettiva. È chiaro che $n > 1$. Senza perdere in generalità possiamo supporre che n non appartenga all'immagine di α. Allora la restrizione di α a $\{1, \dots, n-1\}$ dà una applicazione iniettiva ma non suriettiva di questo insieme in sè, contro la minimalità di n.

Supponiamo ora che esista una famiglia $\mathcal{F}$ di insiemi con la proprietà che ogni insieme è equipotente a uno e un solo insieme in $\mathcal{F}$. Una famiglia con questa proprietà si può costruire canonicamente a partire dagli assiomi della teoria degli insiemi, ma qui non discuteremo questo punto delicato. Chiameremo gli elementi di $\mathcal{F}$ *numeri cardinali*, o semplicemente *cardinali*. La *cardinalità* di un insieme è l'elemento di $\mathcal{F}$ a cui esso è equipotente. Dunque due insiemi hanno la stessa cardinalità se e solo se sono equipotenti. Se α e β sono cardinali, scriveremo $\alpha \leq \beta$

per indicare che α è equipotente a un sottoinsieme di β o, che è lo stesso, che esiste una applicazione suriettiva di β su α. Il fatto, non ovvio a priori, che se $\alpha \leq \beta$ e $\beta \leq \alpha$ allora $\alpha = \beta$, segue dal seguente risultato, che va sotto il nome di teorema di Cantor-Bernstein.

Teorema 5. (Cantor-Bernstein) *Siano A e B due insiemi, e $\alpha : A \to B$, $\beta : B \to A$ due applicazioni iniettive. Allora esiste una applicazione biunivoca da A a B.*

Dimostrazione. Poniamo

$$A_0 = \{x \in A : y \in (\beta\alpha)^n(A) \text{ per ogni } n \in \mathbb{N}\},$$
$$A_+ = \{x \in A : \text{esiste } n \in \mathbb{N} \text{ tale che } x \in (\beta\alpha)^n(A) \text{ ma } x \notin \beta(\alpha\beta)^n(B)\},$$
$$A_- = \{x \in A : \text{esiste } n \in \mathbb{N} \text{ tale che } x \in \beta(\alpha\beta)^n(B) \text{ ma } x \notin (\beta\alpha)^{n+1}(A)\}.$$

e definiamo in modo analogo B_0, B_+ e B_-. È chiaro che A_0, A_+ e A_- formano una partizione di A, e B_0, B_+ e B_- una di B. È anche chiaro che $\alpha(A_0) \subset B_0$, $\alpha(A_+) \subset B_-$ e $\beta(B_+) \subset A_-$. Osserviamo poi che queste inclusioni sono tutte uguaglianze. Supponiamo infatti che $x \in B_0$. Segue dalle definizioni che $x = \alpha(y)$ per qualche y. D'altra parte, per ogni n, possiamo scrivere $x = (\alpha\beta)^{n+1}(z)$ per qualche z. Dunque $\alpha(y) = \alpha(\beta\alpha)^n\beta(z)$. Dato che α è iniettiva ne segue che $y = (\beta\alpha)^n\beta(z) \in (\beta\alpha)^n(A)$. Lo stesso ragionamento mostra che, se $x = \alpha(y)$ e $x \in (\alpha\beta)^n\alpha(A)$, allora $y \in (\beta\alpha)^n(A)$. D'altra parte, se $y \in \beta(\alpha\beta)^n(B)$, allora $x \in (\beta\alpha)^{n+1}(A)$. Dunque $\alpha(A_+) = B_-$, e analogamente $\beta(B_+) = A_-$. In conclusione, l'applicazione $\gamma : A \to B$ definita da

$$\gamma(x) = \begin{cases} \alpha(x) & \text{se } x \in A_0 \cup A_+ \,, \\ \beta^{-1}(x) & \text{se } x \in A_- \,. \end{cases}$$

è biunivoca.

Le cardinalità degli insiemi finiti sono in corrispondenza biunivoca con i numeri naturali. Se n è un numero naturale diremo quindi che un insieme ha cardinalità n se è equipotente a $\{1, \ldots, n\}$. Diremo anche che l'insieme vuoto ha cardinalità 0. Alla cardinalità di $\mathbb{N}$ è riservato un simbolo particolare, e cioè $\aleph_0$ (aleph con zero).

Se α e β sono numeri cardinali, la loro somma $\alpha + \beta$ è definita come la cardinalità di un insieme che sia unione di insiemi disgiunti di cardinalità α e β, mentre il prodotto $\alpha\beta$ è definito come la cardinalità del prodotto cartesiano di un insieme di cardinalità α e di uno di cardinalità β. La potenza α^β è invece la cardinalità dell'insieme di tutte le applicazioni da B in A, dove A ha cardinalità α e B ha cardinalità β; questo insieme viene a volte indicato con A^B. La notazione è giustificata dal fatto che il numero delle applicazioni da $\{1, \ldots, m\}$ a $\{1, \ldots, n\}$ è n^m. Lasciamo al lettore il compito di verificare che tutte queste sono buone definizioni.

Esercizio 4. Perchè non ha molto senso parlare di differenza di numeri cardinali?

Proposizione 2. *Per ogni cardinale β, $2^\beta > \beta$. Più in generale, se $\alpha > 1$, $\alpha^\beta > \beta$.*

È chiaro che basta dimostrare la prima di queste affermazioni. Sia B un insieme di cardinalità β. Ricordiamo che l'insieme $\{0,1\}^B$ è in corrispondenza biunivoca con l'insieme delle parti di B; la corrispondenza è data dall'associare a una applicazione $\varphi : B \to \{0,1\}$ il sottoinsieme $\varphi^{-1}(0)$. È chiaro che $2^\beta \geq \beta$; infatti l'applicazione da B nell'insieme delle parti di B che associa a ogni $b \in B$ il sottoinsieme $\{b\}$ è iniettiva. Bisogna mostrare che $2^\beta \not\leq \beta$, cioè che non ci sono applicazioni suriettive da B sull'insieme delle parti di B. La dimostrazione si basa sul cosiddetto procedimento

diagonale di Cantor. Sia ψ una applicazione da B all'insieme delle parti di B. L'insieme $C = \{x \in B : x \notin \psi(b)\}$, non appartiene all'immagine di ψ. Se infatti si suppone che C sia della forma $\psi(b)$ per qualche b si ottiene una contraddizione. In effetti, se si suppone che $b \in C$, si conclude che $b \notin \psi(b) = C$, per la definizione di C, mentre se si suppone che $b \notin C = \psi(b)$ si conclude invece che $b \in C$, sempre per la definizione di C.

Proposizione 3. $\aleph_0 + \aleph_0 = \aleph_0\aleph_0 = \aleph_0$.

L'insieme $\mathbb{N}'$ dei naturali pari è equipotente a $\mathbb{N}$; una applicazione biunivoca dal secondo insieme al primo è data da $n \mapsto 2n$. Anche l'insieme $\mathbb{N}''$ dei naturali dispari è equipotente a $\mathbb{N}$; una applicazione biunivoca dal secondo insieme al primo è data da $n \mapsto 2n - 1$. D'altra parte $\mathbb{N}$ è unione disgiunta di $\mathbb{N}'$ e $\mathbb{N}''$; questo mostra che $\aleph_0 + \aleph_0 = \aleph_0$. Consideriamo ora l'applicazione da $\mathbb{N} \times \mathbb{N}$ in $\mathbb{N}$ data da $(n, m) \mapsto 2^n 3^m$. Questa applicazione è iniettiva per la proprietà di fattorizzabilità unica degli interi, quindi $\aleph_0\aleph_0 \leq \aleph_0$. D'altra parte vale anche la disuguaglianza opposta, ad esempio perchè l'applicazione da $\mathbb{N}$ in $\mathbb{N} \times \mathbb{N}$ data da $n \mapsto (n, 1)$ è iniettiva. Segue dal teorema di Cantor-Bernstein che $\aleph_0\aleph_0 = \aleph_0$.