L'impostazione categoriale della matematica*

Alberto Peruzzi

Dipartimento di Filosofia, Università di Firenze

1. Considerazioni preliminari

Nel Novecento c'è stata una crescita, senza pari nella storia, di nuovi settori dell'indagine scientifica. A questa crescita si è accompagnato un corrispondente incremento di curricoli formativi sempre più specializzati. Al contempo, il Novecento ha visto nascere e svilupparsi straordinari progetti di unificazione del sapere. L'esigenza di dar conto di fatti anche molto particolari, ben *all'interno* di un ambito di ricerca, può far avvertire la necessità di un più nitido assetto dei fondamenti teorici; e, quando ci si impegna in questo compito, il lavoro di ricerca conduce non di rado a sentire l'esigenza di un linguaggio in grado di esprimere *relazioni sistematiche* fra le conoscenze relative all'ambito in questione e quelle relative ad altri ambiti della stessa disciplina. Si creano allora le condizioni per un tentativo di unificazione del sapere, che può propagarsi in ambiti inizialmente non intesi e poi in discipline diverse.

Ciò che contraddistingue alcuni dei principali progetti di unificazione coltivati nella scienza del Novecento è il fatto che essi hanno dato alla riflessione (meta-teorica) sui fondamenti di una disciplina uno status rigoroso e ne hanno fatto un oggetto di studio alla pari degli oggetti concretamente indagati nella disciplina stessa. Per esempio, in fisica si studiano i principi d'invarianza delle leggi fisiche. Accanto ai progetti di unificazione che non hanno ancora avuto un chiaro esito o che sono stati abbandonati, ci sono quelli che hanno avuto un successo tale da farne un elemento portante della nostra attuale immagine del mondo, producendo significative trasformazioni anche nella didattica di numerose discipline scientifiche.

In biologia, la scoperta del DNA ha permesso di fornire una descrizione unitaria dei mattoni fondamentali della vita, a partire dai quali le differenze tra le varie specie sono riconducibili a sequenze diverse in uno stesso codice. In fisica, a partire dalla teoria relativistica della gravità e dalla meccanica quantistica, si è avviata la ricerca di una grande unificazione che, determinando i principi di correlazione fra le quattro forze fondamentali, integri in un quadro coerente microfisica e cosmologia. In linguistica, si è giunti a una teoria generale delle grammatiche come algoritmi di produzione di frasi, in grado di identificare la struttura di ogni lingua naturale.

In matematica, i vari ambiti tradizionali (teoria dei numeri, algebra, analisi e geometria) hanno trovato espressione unitaria prima nel linguaggio della teoria degli insiemi e poi nel linguaggio della teoria delle categorie. La matematica è strumento essenziale della fisica; la fisica fa da sfondo alla chimica; e la biologia molecolare odierna s'impernia sulla chimica organica; ma la matematica è anche strumento essenziale della teoria linguistica. Il modo in cui la matematica trova la sua unificazione non è privo di conseguenze sulle applicazioni che la matematica ha in settori già consolidati delle scienze o in settori di frontiera. Per esempio, le basi dell'informatica non sono indifferenti alla scelta di un'impostazione insiemistica o categoriale.

Ci sono subito due considerazioni da fare, in corrispondenza con due diffusi frantendimenti. Uno riguarda la natura della matematica e un altro la funzione che in essa svolgono teorie come quella degli insiemi e delle categorie.

Quanto al primo, si sente dire spesso che la matematica è un linguaggio, che è poi il linguaggio canonico della scienza. Il che non è semplicemente falso: è assurdo. Un linguaggio non fornisce di per sé conoscenze, ma soltanto un mezzo per esprimerle. Dunque non esisterebbe un sapere matematico, mentre invece esiste. E si articola in teoremi su procedure di calcolo, strutture algebriche, spazi ... Chi suppone, altresì, che la matematica esprima un sapere esclusivamente su linguaggi simbolici, introduce un uso improprio (che direi metafisico) del termine "linguaggio" oppure confonde numeri, derivate, gruppi e superfici con i simboli che usiamo per descriverli. Infine, non va dimenticato che esistono teorie matematiche alternative: ci sono geometrie diverse, formulabili nello stesso linguaggio, che si diversificano per i teoremi in esse dimostrabili; esiste un'analisi standard (con proprietà archimedea del campo dei reali, definizione ε–δ di continuità, ecc.) e una non-standard, ci sono varie teorie degli insiemi e perfino logiche diverse, che concordano su alcuni risultati e divergono su altri. Per la maggior parte degli usi pratici, non è necessario tenere conto di queste differenze, ma la molteplicità di sistemi cui esse conducono è, oltre che una delle caratteristiche primarie della matematica del Novecento, un oggetto di studio della matematica stessa.

Il secondo fraintendimento è dovuto all'idea che la teoria degli insiemi fornisca, se non *il* linguaggio, *l'unico* quadro teorico in cui formulare, con il rigore richiesto, le nozioni e i principi basilari delle varie branche della matematica. Ora, anche a prescindere dal fatto già ricordato che ci sono più teorie degli insiemi (e, visto il loro sviluppo, si può ormai considerare 'la' teoria degli insiemi come una vera e propria branca fra le altre della matematica), il linguaggio insiemistico e i principi in esso esprimibili rappresentano una scelta tutt'altro che scontata. Il loro impiego corrente nella didattica, a partire dall'*insiemistica*, oltre a non tener conto delle reali motivazioni che giustificano l'introduzione di questo linguaggio, lascia supporre che non ci sia altro modo di presentare le nozioni fondamentali della matematica. Il che è falso, perché la teoria delle categorie offre appunto un modo diverso e non meno efficace, anche se sul piano didattico l'opportunità di impiegare l'una o l'altra può dipendere dall'argomento trattato. (Passando sotto silenzio le motivazioni, tanto varrebbe continuare a usare la nozione di insieme come intuitiva e tornare a insegnare analisi e geometria come in passato.)

È vero che anche in matematica, così come in fisica, biologia e linguistica, per limitarci alle aree su menzionate, ci sono molte cose che si possono dire sensatamente, e sul piano didattico forse in maniera più efficiente, senza chiamare in causa i principi primi, cioè, i fondamenti teorici a partire dei quali ricostruiamo, nella maniera più generale, le conoscenze circa la nostra ordinaria esperienza di fatti (fenomeni, dati, sistemi) fisici, biologici, linguistici ... e appunto matematici. Ciononostante, già in relazione a problemi matematici elementari si può trarre vantaggio orientando l'attenzione su nozioni e procedure ricorrenti da un problema all'altro, se si dispone di un quadro teorico di riferimento che unisca flessibilità, facilità d'uso in nuovi casi (trasferibilità) e profondità concettuale.

La teoria degli insiemi e la teoria delle categorie sono due straordinarie conquiste del pensiero umano. Entrambe combinano semplicità dei concetti di base e potenza dei loro sviluppi, entrambe offrono un linguaggio *universale* e un corrispondente quadro teorico di riferimento, entrambe si propongono di unificare tutta la matematica; hanno però principi, motivazioni e modalità costruttive alquanto diversi. Anche se qui non è possibile entrare nel dettaglio delle loro differenze, né soppesare i rispettivi vantaggi e svantaggi, è opportuno richiamare l'attenzione su alcuni aspetti generali del confronto, perché la stessa didattica della matematica ne è toccata – e in maniera sostanziale.

Purtroppo, l'esistenza della teoria delle categorie non è molto nota e, comunque, si tende ad assimilarla a una specie di algebra generalizzata la cui utilità resta appunto confinata all'algebra. In tal modo diventa arduo capire come la teoria delle categorie permetta allo stesso tempo di introdurre una nozione più generale di insieme, reimpostare i principi della logica, fondare l'analisi, descrivere un concetto di spazio di cui quello accessibile mediante la topologia degli insiemi di punti è solo un caso particolare, e fornire una cornice appropriata per la semantica dei linguaggi di programmazione.

Allorché si riduce il confronto fra insiemi e categorie a quello fra linguaggi alternativi, si cade nel primo fraintendimento su considerato. L'impostazione categoriale non porta agli stessi risultati ottenibili servendosi della teoria degli insiemi: è falso che tutto quanto è esprimibile e dimostrabile mediante metodi categoriali lo sia già nel quadro insiemistico standard.

C'è un ultimo preliminare da rendere esplicito – ed è di carattere propriamente "pedagogico". Tanto la teoria degli insiemi quanto la teoria delle categorie offrono una sistemazione generale del pensiero matematico in un disegno unitario, ma questo non implica che si debbano introdurre i giovani alla matematica *partendo* dai fondamenti che l'una o l'altra teoria fornisce alla matematica nel suo complesso. Ciò, per il semplice motivo che, al fine di apprezzare il significato del rispettivo disegno fondazionale, occorre già sapere un bel po' di matematica – è un'osservazione di Lolli che condivido pienamente e alla quale mi limito ad aggiungere che proprio nell'impostare problemi concreti di algebra e di geometria si capisce il senso … e la bellezza dei principi.

Tuttavia, nessun linguaggio nasce sotto vuoto; e usarlo *come se* fosse sotto vuoto produce più danni che benefici. Questo è specialmente riconoscibile se si tiene conto che la matematica è da sempre il frutto di sistematiche correlazioni fra entità e costrutti di tipo diverso: numeri, figure, algoritmi, simmetrie … È proprio a tale proposito che la teoria delle categorie mostra un significativo vantaggio sulla teoria degli insiemi, perché orienta a pensare in termini di rapporti dinamici fra strutture invece che statici, in termini uniformi e invarianti invece che caso per caso, e a recuperare, per quanto ciò di primo acchito possa suscitare incredulità, le capacità intuitive di organizzazione razionale dei pensieri, attraverso la valorizzazione dell'aspetto diagrammatico e procedurale.

Della teoria degli insiemi, in versione intuitiva o in versione assiomatica, si possono trovare molte, e ben fatte, esposizioni in italiano, unitamente a testi che ora ne ricostruiscono lo sviluppo storico, ora ne affrontano la problematica filosofica, ora ne discutono le modalità di presentazione sul piano didattico. Al contrario, sulla teoria delle categorie ci sono pochissime pubblicazioni, tanto meno accessibili a un pubblico di non specialisti. Fra quelle esistenti, la più elementare è costituita da [4], che però non fornisce un'introduzione storica; limitatamente all'algebra, l'impiego sistematico delle categorie è documentato in [5]; mentre l'unica esposizione che ricostruisce alcuni momenti dello sviluppo storico della teoria è fornita da [6], ma è orientata a metterne in evidenza gli aspetti di rilievo per la logica. Cercherò dunque di dare un'idea, in forma estremamente succinta, delle motivazioni che stanno all'origine della teoria delle categorie (§2), passando poi a indicarne i concetti centrali e a menzionare alcuni risultati significativi (§3); accennerò infine ai problemi concernenti il rapporto tra prospettiva categoriale e insiemistica (§4).

2. Motivazioni e sviluppi

Nel 1945 Samuel Eilenberg e Saunders Mac Lane pubblicarono un articolo che è considerato il lavoro inaugurale della teoria delle categorie. Al concetto di categoria si era giunti come al distillato di una serie di passaggi da nozioni complesse a ciò che era richiesto per definirle in modo corretto: dalla nozione di trasformazione naturale a quella di funtore e da questa, infine, alla nozione di categoria.

Consideriamo uno spazio vettoriale V di dimensione finita sul campo dei reali $\mathbf{R}$. L'insieme V^* delle applicazioni lineari a valori in $\mathbf{R}$ è anch'esso uno spazio vettoriale, con $(f \oplus g)(v) = f(v) + g(v)$ e $(\lambda f)(v) = \lambda \cdot f(v)$ per ogni vettore v di V e ogni scalare λ in $\mathbf{R}$. V^* si chiama *spazio duale* di V. Dato che la dimensione di V è finita, V^* risulta isomorfo a V, ma non c'è alcun modo 'naturale' (canonico) di definire questo isomorfismo, cioè un modo che sia indipendente dalla base. Tuttavia, per il duale di V^*, cioè il *doppio duale* di V, questo modo c'è. Per definire l'isomorfismo (canonico) $V \longrightarrow V^{**}$ basta associare a ogni $v \in V$ l'applicazione $v^{\sim}$: $V^* \rightarrow \mathbf{R}$ t. c. $v^{\sim}(f) = f(v)$ che è biiettiva e lineare; così facendo, si trasforma l'applicazione lineare identica su V ($v \mapsto v$) nell'applicazione che associa a V il doppio duale V^{**} – e ciò in modo naturale.

Ma come si può dire, in generale, cos'è una trasformazione «naturale»? Per rispondere bisogna prima aver precisato alcuni concetti implicitamente usati. Vediamo quali, attraverso un esempio paradigmatico. Infatti, all'origine della teoria c'era anche un motivo più sostanzioso, dovuto all'esigenza di descrivere la sistematica correlazione fra spazi e gruppi che è al centro della topologia algebrica. Questa correlazione si precisa in vari modi, ma forse quello più intuitivo passa per l'associazione a ogni spazio del suo gruppo fondamentale di omotopia.

Dato uno spazio X, un cammino in X è una funzione continua $f: I \rightarrow X$, ove I è l'intervallo chiuso $[0,1] \subset \mathbf{R}$. Il punto iniziale del cammino sarà il punto a di X tale che $a = f(0)$ e quello finale sarà il punto b di X tale che $b = f(1)$. Quando $f(0)=f(1)$, il cammino è chiuso. Due cammini f e g si possono comporre quando il punto finale del primo è il punto iniziale del secondo. Il cammino nullo è quello costante. Quando due cammini f, f' che per semplicità abbiano lo stesso punto iniziale $f(0) = f'(0)$ e finale $f(1) = f'(1)$ sono deformabili con continuità l'uno nell'altro si dicono omotopi. Quella di omotopia è una relazione di equivalenza e ripartisce appunto i cammini in classi di equivalenza. Supponiamo che lo spazio X sia connesso, cioè, non possa essere diviso in due aperti non vuoti disgiunti U, V tali che $U \cup V = X$. Allora, una volta fissato un arbitrario punto base x_0 in X l'insieme dei cammini chiusi su x_0, ciascuno dei quali è identificato a meno di tale equivalenza, forma un gruppo che non dipende da x_0 e che si chiama (primo) gruppo (fondamentale) di omotopia $\pi_1(X)$. Per esempio, il gruppo di omotopia di una sfera è diverso da quello di un toro.

Ora si tratta di mettere insieme i fatti seguenti: spazi omeomorfi $X \cong Y$ hanno gruppi di omotopia isomorfi $\pi_1(X) \cong \pi_1(Y)$; ogni funzione continua fra spazi $f: X \rightarrow Y$ induce un omomorfismo f^* di gruppi di omotopia; se $g: Y \rightarrow Z$ è un'altra funzione continua, la composta $gf: X \rightarrow Z$ è continua e $(gf)^* = (g)^*(f)^*$; la composizione di funzioni continue è associativa e tale è pure quella di omomorfismi di gruppi. Nel caso che $Z = X$, $g = f^{-1}$ e biiettiva, si ha $X \cong Y$ e $(f^{-1})^*(f)^* = (f^{-1}{}^*f^*)$ è l'identità su $\pi_1(X)$, mentre $(f)^*(f^{-1})^*$ è l'identità su $\pi_1(Y)$, ovvero f^* è un isomorfismo di gruppi.

Dobbiamo trovare un nome alla corrispondenza generale '*' tra spazi e gruppi. Per un oggetto X arbitrario di un dato tipo, indicheremo l'identità su X con 1_X. Diremo che ogni corrispondenza F tra oggetti di un tipo e oggetti di un altro che conservi l'identità, cioè $F(1_X) = 1_{FX}$, e che conservi la composizione $F(gf) = F(g)F(f)$ è un *funtore* (covariante). Ma come è già chiaro dall'esempio, non stiamo considerando soltanto due tipi di 'cose' – tutti gli oggetti di un tipo (spazi) e tutti quelli di un altro (gruppi) –, ma anche due modi di correlare tra loro gli oggetti: il modo di correlare tra loro oggetti del primo tipo e il modo di correlare tra loro quelli del secondo tipo, dunque non solo spazi e gruppi, ma anche funzioni continue tra spazi e omomorfismi tra gruppi. Allora, che cosa devono essere queste due totalità che un funtore collega? Devono essere … *categorie*.

Una categoria $\mathbf{C}$ è costituita da *oggetti* A, B, C, ..., e da *morfismi* f, g, h, … tra oggetti, ove ciascun morfismo $f: A \rightarrow B$ ha A come *dominio* e B come *codominio* – si scrive solitamente $A \xrightarrow{f} B$ e, dati $A \xrightarrow{f} B$ e $B \xrightarrow{g} C$, è definito un morfismo $A \xrightarrow{g \circ f} C$ che è la composizione di f e g (il simbolo $\circ$ è solitamente omesso). Inoltre, per ogni oggetto X c'è un morfismo *identità* (sull'oggetto dato), che come sopra si scrive

1_X, ed è caratterizzato dal fatto che, per qualunque f da un arbitrario Y a X, $1_X f = f = f\,1_Y$. Infine, la composizione gode della proprietà associativa, cioè $h(gf) = (hg)f$, purché le varie composizioni siano definite.

La categoria **Set** degli insiemi è formata da oggetti che sono, appunto, gli insiemi e da morfismi che sono le usuali funzioni: l'identità è la funzione che associa a ogni elemento a di un insieme X lo stesso a, $1_X(a) = a$ e le funzioni si compongono in modo ovvio. La categoria **Top** degli spazi topologici ha come oggetti spazi e come morfismi funzioni continue. La categoria **Grp** dei gruppi ha come oggetti i gruppi e come morfismi gli 'omomorfismi'. Un altro esempio è fornito dalla categoria degli ordini lineari in cui un oggetto è un qualsiasi ordine lineare di punti e un generico morfismo f da un ordine lineare A a un altro B è una corrispondenza univoca da A a B che conserva la relazione di "venire prima di" tra punti: se p e q sono punti di X e $p \leq q$ in A allora $f(p) \leq f(q)$ in B. Si noti che $\leq$ non è una funzione. La categoria **Prop** delle proposizioni ha come oggetti proposizioni $\phi, \psi, \ldots$, e come morfismi $\phi \to \psi$ le prove di ψ a partire da ϕ.

Date due categorie **C** e **D**, la totalità dei funtori da **C** a **D** forma una categoria, prendendo come morfismi $\tau\colon \boldsymbol{F} \to \boldsymbol{G}$ tra due funtori da **C** a **D** le *trasformazioni naturali*, definite in modo che indipendentemente da quali siano gli oggetti X, Y di **C** e i morfismi h tra essi, il diagramma seguente risulti commutativo, cioè: $G(h) \circ \tau_X = \tau_Y \circ F(h)$.

$$\begin{tikzcd} \boldsymbol{C} & & \boldsymbol{D} & \\ X \arrow[d, "h"'] & & F(X) \arrow[r, "\tau_X"] \arrow[d, "F(h)"'] & G(X) \arrow[d, "G(h)"] \\ Y & & F(Y) \arrow[r, "\tau_Y"] & G(Y) \end{tikzcd}$$

Se consideriamo le categorie come oggetti e i funtori come morfismi tra essi, otteniamo la categoria delle categorie: **Cat**. Come l'insieme V di tutti gli insiemi dava luogo a un paradosso, perché (Teorema di Cantor) l'insieme potenza $\mathscr{P}(V)$ dovrebbe essere più grande di V, anche in questo caso sorgono problemi che si evitano limitando la 'taglia' delle categorie che possono essere oggetti in **Cat**.

Tutte queste definizioni non possono dare un'idea delle costruzioni categoriali che, con elegante uniformità, consentono di caratterizzare in maniera appropriata prodotti, quozienti e vari altri modi di ottenere nuove strutture, algebriche e topologiche, a partire da strutture date. Nondimeno, è già possibile rendersi conto che il quadro risultante è guidato dalla preoccupazione di fornire un linguaggio e una teoria in grado di cogliere le proprietà che legano fra loro i vari tipi di nozioni matematiche, con l'intento di mettere in risalto un'architettura modulare il più possibile aderente a ciò che i matematici fanno, quando costruiscono strutture per poi analizzarne i componenti (si pensi alla fattorizzazione di un gruppo in sottogruppi) e indagarne le proprietà mappando strutture di un tipo in quelle di un altro (si pensi alla misurazione).

Per più di un decennio la teoria delle categorie restò confinata all'algebra e alle sue diramazioni. Alla fine degli anni cinquanta, si verificarono due scoperte decisive. Da un lato, fu introdotto un concetto che poi è risultato quello centrale per lo sviluppo della teoria delle categorie e per le sue applicazioni alla logica e all'informatica teorica: il concetto di *aggiunzione* fra funtori. Dall'altro, il matematico francese Alexandre Grothendieck si servì delle categorie per risolvere una serie di problemi di geometria algebrica e, nel fare questo, introdusse un altro concetto chiave: il concetto di *topos*, come spazio generalizzato, non più dipendente dalla nozione di punto, ma dato in termini di ricoprimenti sugli oggetti di una opportuna categoria (*sito*).

Nel corso degli anni sessanta, il matematico americano Bill Lawvere riuscì a capire come si potevano descrivere in linguaggio puramente categoriale, senza fare uso del concetto di appartenenza (∈) le proprietà degli insiemi; insieme a Myles Tierney introdusse nel 1971 il concetto di topos elementare e riuscì a mostrare come tutti i concetti della logica si potevano ottenere in termini di funtori aggiunti. (Insieme al concetto di aggiunzione anche quello di topos elementare sarà definito nel §3.) L'idea stessa di fondamenti della matematica ne veniva profondamente modificata: gli assiomi posti come principi fondamentali devono corrispondere ai principi universali che guidano la pratica matematica nei suoi più diversi ambiti, e non semplicemente un insieme di principi grazie ai quali si possa garantire l'esistenza di opportuni insiemi (assioma dell'infinito, assioma di scelta). Si può dire che il periodo dal 1958 al 1971 è stato il periodo eroico della teoria delle categorie, cui è seguito un periodo di espansione verso nuovi settori di ricerca, di sempre più accurata formulazione, nonché di sistemazione organica dei risultati conseguiti. Quello che sembrava un *nonsense* astratto, in cui non si capisce di che cosa si parla e in cui non si vede quali teoremi dimostrare che non si dimostrassero già, è diventato un potente strumento euristico che ha permesso di ottenere risultati non solo di notevole importanza ma neppure lontanamente immaginabili prima del suo impiego.

3. Concetti universali

Tanto per cominciare, una nozione basilare della matematica moderna, che la teoria delle categorie consente di definire in modo *indipendente dalla particolare categoria considerata*, è quella di *isomorfismo*: due oggetti A e B di una categoria sono isomorfi quando esiste un morfismo $f\colon A \to B$ che sia invertibile, cioè per il quale c'è un $g\colon B \to A$ t. c. $gf = 1_A$ e $fg = 1_B$. Quindi $g = f^{-1}$ e $f = g^{-1}$.

Inoltre, tutte le nozioni insiemistiche sono riottenibili in forma puramente categoriale: per esempio, il fatto che una funzione f da un insieme A a un insieme B sia *iniettiva* si esprime dicendo che, per ogni *h*, *k* da qualsiasi *X* ad *A*, $fh = fk$ implica $h = k$. Questa proprietà categoriale definisce un *monomorfismo.* L'inclusione di un sottoinsieme in un insieme è una particolare funzione iniettiva. In termini di monomorfismi si definisce in una categoria arbitraria **C** il concetto di *sottoggetto*, come classe di equivalenza di monomorfismi: un monomorfismo $U \to A$ è equivalente a un altro $V \to A$ quando *U* e *V* sono isomorfi in **C**. La nozione di sottoggetto è intrinseca, quella di sottoinsieme no. Per esempio, in **Grp** i sottoggetti non sono meri sottoinsiemi ma sottogruppi, in **Top** i sottoggetti devono essere sottospazi, quindi devono ereditare una topologia dallo spazio ambiente. La nozione duale di monomorfismo è quella di *epimorfismo*: per ogni *h*, *k* da *A* a un qualsiasi *X*, $hf = kf$ implica $h = k$. In **Set** questa nozione coincide con quella di funzione suriettiva; in altre categorie non è così.

Passiamo alle modalità di costruzione di nuovi oggetti. Il concetto di prodotto cartesiano è usualmente definito facendo riferimento a coppie (n-ple) di elementi dei fattori, ma in linguaggio categoriale il prodotto può essere dato senza fare riferimento agli elementi e in modo tale da rendere la definizione indipendente dalla particolare categoria ambiente (si tratti del prodotto di gruppi o di uno spazio prodotto, per esempio). Una categoria **C** ha prodotti se per ogni due oggetti *A, B* esiste un oggetto $A \times B$ con due morfismi 'di proiezione' p_1: $A \times B \to A$ e p_2: $A \times B \to B$, tali che per ogni oggetto *X* e ogni f: $X \to A$, g: $X \to B$, esiste un unico h: $X \to A \times B$ che rende *commutativo* il seguente diagramma, cioè $p_1 h = f$ e $p_2 h = g$.

$$\begin{tikzcd} & X \arrow[dl, "f"'] \arrow[d, dashed, "h"] \arrow[dr, "g"] & \\ A & A \times B \arrow[l, "p_1"] \arrow[r, "p_2"'] & B \end{tikzcd}$$

L'oggetto così definito è unico a meno di isomorfismi: in effetti, $A \times B \cong B \times A$. Nel caso di **Grp** questa costruzione universale dà il prodotto di gruppi e nel caso di **Top** dà lo spazio prodotto. Inoltre, se la categoria è un insieme parzialmente ordinato dall'inclusione, l'esistenza del prodotto di due oggetti esprime il fatto che esiste la loro intersezione; mentre nella categoria **Prop** l'esistenza del prodotto così definito esprime le proprietà logiche che caratterizzano la congiunzione. Si noti che la *definizione* di prodotto su illustrata non sarebbe possibile nella teoria degli insiemi, ove i morfismi (cioè, le funzioni) sono introdotti come relazioni univoche – cioè le *R* tali che $R(x,y)$ e $R(x,z)$ implica $y = z$ –, perché le relazioni sono sottoinsiemi di un prodotto cartesiano, che dunque deve essere dato prima del concetto di funzione e non dopo.

In una categoria con prodotti, dati f: $A \to C$ e g: $B \to D$, risulta definito anche il morfismo prodotto $f \times g$: $A \times B \to C \times D$. La definizione di prodotto si riferisce a due oggetti, ma si estende facilmente a un arbitrario numero finito di oggetti: si può dimostrare che $(A \times B) \times C \cong A \times (B \times C)$. Nel caso che questo numero finito sia zero, si ha come caso particolare il semplice fatto che esiste un oggetto 1 tale che per ogni X c'è uno e un solo morfismo $X \to 1$. Tale oggetto, che si chiama *oggetto terminale* di **C** è unico a meno di isomorfismi; il morfismo unico $X \to 1$ si indica con $!_X$. Nella categoria degli spazi, 1 è lo spazio con un solo punto; in **Grp** è il gruppo banale con il solo elemento neutro. In **Set** è un qualsiasi singoletto $\{\bullet\}$. Dato che c'è una corrispondenza biunivoca fra A e i morfismi (funzioni) $\{\bullet\} \to A$, ciascuno di questi può essere pensato come un *elemento* di A.

Il duale della nozione di prodotto è quella di coprodotto (o somma) e si ottiene invertendo tutti i morfismi nella figura precedente – le proiezioni diventano iniezioni degli addendi nella somma. Nella categoria degli insiemi il coprodotto di due insiemi è la loro unione disgiunta (dunque questa viene prima, per ragioni strutturali, della semplice unione: gli elementi comuni non possono essere contati due volte). Il duale dell'oggetto terminale è l'oggetto *iniziale*: una categoria **C** ha oggetto iniziale se esiste un oggetto 0 tale che esiste uno e un solo morfismo $0 \to X$, per ogni X in **C**. In **Grp** è di nuovo il gruppo banale, in **Set** è … l'insieme vuoto.

Fondamentali in matematica sono le categorie che oltre ad avere prodotti (finiti) sono *chiuse*, nel senso che la collezione dei morfismi da un oggetto A a un oggetto B è a sua volta un oggetto B^A ('spazio di funzioni' da A a B) che è individuato dalla seguente proprietà universale: per ogni $f\colon A \times B \to C$, esiste un unico morfismo $f^\wedge\colon B \to C^A$ che rende commutativo il diagramma seguente, ove il morfismo ev_{AC} non è altro che la valutazione in C di un morfismo f per un argomento in A. Cioè: $f = ev_{AC} \cdot (1_A \times f^\wedge)$.

$$\begin{tikzcd}
A \times C^A \arrow[dr, "ev_{AC}"] & \\
A \times B \arrow[u, dotted, "1_A \times f^\wedge"] \arrow[r, "f"'] & C
\end{tikzcd}$$

Le categorie che oltre ad avere prodotti soddisfano questa proprietà si dicono *cartesiane chiuse*. In informatica teorica, il passaggio da f a $f^\wedge$ è di grande utilità perché serve a trasformare in parametri le variabili aggiuntive di una funzione di più argomenti (in figura solo due). Nel caso di un dominio X formato da procedure di calcolo che dovessero essere sempre applicabili l'una all'altra senza uscire dal dominio, bisognerebbe che $X^X \cong X$. Un tale X non può essere un insieme, perché l'unico insieme tale che le funzioni da X in X non siano più degli elementi di X è un singoletto. Esistono invece categorie non banali che permettono questa rappresentabilità interna delle procedure di calcolo.

Come ci si può immaginare, anche il modo in cui si dimostra qualcosa in teoria delle categorie è un po' diverso da quello divenuto consueto con l'uso della teoria degli insiemi. In particolare, invece di sfruttare ragionamenti per induzione, si sfruttano le proprietà che contraddistinguono un oggetto a meno di isomorfismo e le costruzioni universali che sono possibili nella categoria data, espresse in forma diagrammatica, come nel caso del prodotto. Questa forma corrisponde alla priorità accordata a un linguaggio equazionale, che è appunto il tipo di linguaggio in cui si esprime al meglio la nostra conoscenza della natura.

I rapporti con la logica vanno ben oltre l'esempio della congiunzione. Anche gli altri connettivi sono definibili in forma diagrammatica e, soprattutto, le proprietà che definiscono i quantificatori (universale $\forall$ ed esistenziale $\exists$) sono riottenibili mediante equazioni. A questo scopo è stato decisivo il concetto di aggiunzione, che in realtà è il concetto centrale di tutta quanta la teoria delle categorie. La nozione di prodotto e le altre che ho introdotto a partire da quelle di categoria e di funtore, fino a coprire tutte le costruzioni che ricorrono da una teoria matematica all'altra, sono infatti esprimibili come aggiunzioni. Vediamo allora come si definisce un'aggiunzione (un primo esempio fu identificato da Galois, con la 'corrispondenza' che reca il suo nome in teoria dei campi).

Un funtore F: $\mathbf{C} \longrightarrow \mathbf{D}$ si dice AGGIUNTO sinistro di un funtore G: $\mathbf{D} \longrightarrow \mathbf{C}$ quando c'è un isomorfismo ϕ tra $\{g: X \longrightarrow GY\}$ e $\{f: FX \longrightarrow Y\}$ che è *naturale*, nel senso che, per ogni $Fh: FX' \rightarrow FX$ e per ogni $Gk: GY \longrightarrow GY'$, la corrispondenza isomorfa si estende da quella tra g a f a quella da gGk a fFh. Se F è aggiunto sinistro di G, G è aggiunto destro di F. Ebbene, per quanto vi possa sembrare incredibile, anche tutte le nozioni logiche si possono descrivere come aggiunti; più in generale, situazioni di aggiunzione si ritrovano ogniqualvolta abbiamo a che fare con costruzioni di concetti che individuano strutture «universali».

Per esempio, ogni volta che si fissano i generatori di uno specifico tipo di struttura, la costruzione di una struttura libera sui dati generatori è un caso particolare di aggiunzione (un caso istruttivo è quello del gruppo libero). Si può infatti dimostrare che tutte le nozioni categoriali precedentemente introdotte (prodotti, coprodotti, ecc.) sono casi particolari di aggiunzione tra funtori. In più, i funtori aggiunti svolgono in logica una funzione decisiva quando si studiano i modelli di una teoria (cioè le strutture in cui la teoria può essere interpretata e in cui i suoi assiomi risultano veri) perché garantiscono la preservazione di formule che esprimono proprietà 'geometriche'.

Sempre per quanto riguarda la logica, l'impostazione categoriale ha anzi consentito di riprendere e sviluppare in modo nuovo la *teoria dei tipi*, con particolare riferimento ai suoi impieghi nell'informatica, ove un *tipo-di-dati* si presenta come un oggetto che può consistere a sua volta di procedure da un tipo-di-dati a un altro. Data una categoria di tipi-di-dati e una base di operazioni primitive «tipate» e alcuni «costruttori», si producono nuovi tipi e nuove operazioni, più complesse, tutte espresse da costruzioni categoriali. La semantica per i linguaggi di programmazione funzionale 'a oggetti' è essenzialmente fornita mediante categorie opportune, fornendo vantaggi analoghi all'analisi dimensionale in fisica.

Per tornare agli ambiti classici della matematica, è possibile assiomatizzare l'aritmetica in termini puramente categoriali, senza far uso di $\in$ e con un solo assioma invece dei cinque usuali nella formulazione di Dedekind-Peano, cfr. [4], ed è possibile costruire successivamente gli altri sistemi di numeri. Tuttavia, si scopre che il continuo dei numeri reali costruito in termini di successioni convergenti e il continuo costruito mediante sezioni di Dedekind non coincidono più come nella matematica classica. Quanto all'analisi, è possibile svilupparla recuperando un'idea originaria: quella secondo cui gli infinitesimi d sono 'elementi' nilpotenti $d^2=0$. Per un'introduzione particolarmente elegante e didatticamente efficace, che si sforza di ridurre al minimo l'apparato categoriale da cui è scaturita, si veda [1]. L'ambiente ideale in cui sviluppare l'analisi è fornito da un topos di tipo particolare, in cui tutte le funzioni risultano lisce, cioè infinitamente differenziabili e a fortiori continue. Con questi strumenti a disposizione, il programma di rifondare la dinamica in termini categoriali è stato avanzato da Lawvere ed è tuttora in corso di elaborazione.

C'è un aspetto importante che contraddistingue alcune delle differenze tra i risultati ottenibili con la teoria delle categorie e quelli ottenibili con la teoria degli insiemi. Nel corso del Novecento sono stati proposti diversi tipi di matematica costruttiva, molti dei quali convergono sul rifiuto del principio logico del terzo escluso e quindi delle dimostrazioni per assurdo, ma anche di numerose definizioni date per distinzione dei casi. Uno dei principali orientamenti è stato quello che va sotto il nome di *matematica intuizionistica*. Le sue motivazioni sono state essenzialmente di tipo epistemico: non si può asserire l'esistenza di qualcosa che non si è in grado di costruire e non si può asserire di aver provato una disgiunzione 'o p o q' se non si è in grado di provare p o di provare q.

Ora, la logica che risulta dalla struttura stessa di un topos è appunto intuizionistica, ma non per tali motivazioni, bensì per le proprietà categoriali che lo definiscono. La matematica classica, anche una volta riformulata categorialmente, diventa un caso limite di una matematica in cui in primo piano sono totalità di oggetti variabili (non rigidamente fissati una volta per tutte) e coesivi (non polverizzati nei loro elementi), su cui si ragiona in modo costruttivo. Senza sviluppare questo punto come meriterebbe, è opportuno notare che nell'impostazione categoriale, accanto al processo di ulteriore algebrizzazione che l'accompagna, un aspetto emergente è anche il recupero del ruolo del pensiero geometrico, che torna finalmente a occupare una posizione centrale e acquista priorità anche nei confronti della stessa logica, cfr. [7].

Per arrivare alla formulazione categoriale del concetto astratto di insieme, dobbiamo prima introdurre un ultimo concetto. Quando si considerano insiemi astratti, una loro proprietà fondamentale è che, per ciascuno di essi, diciamo A, si può formare l'insieme delle parti (o sottoinsiemi) $\mathscr{P}(A)$. A ogni parte $U \subseteq A$ corrisponde una funzione 'caratteristica' $\chi_U\colon A \to \{0,1\}$, ove $\chi_U(x) = 1$ se $x \in U$ e $\chi_U(x) = 0$, se $x \notin U$, per ogni x in A. E viceversa: a ogni funzione $f\colon A \to \{0,1\}$ corrisponde un sottoinsieme di A. Entrambe le corrispondenze sono univoche, perciò $\mathscr{P}(A) \cong \{0,1\}^A$. Qualunque formula esprima una proprietà relativa ad A individua il sottoinsieme degli elementi di A che soddisfano alla proprietà così espressa. Per esempio, la formula $\phi(x)$ che afferma $\exists y\ (x = 2y)$ interpretata sugli interi positivi individua il sottoinsieme dei numeri pari. Allora possiamo, in questo contesto, pensare i due elementi 1 e 0 di {0,1} come rispettivamente il Vero (V) e il Falso (F). Per evitare confusioni, in quanto segue conviene appunto sostituire {V,F} a {0,1}.

Indichiamo con *true* il morfismo $\{\bullet\} \to \{V,F\}$ che a $\bullet$ associa V. Quel che abbiamo appena notato sulla corrispondenza fra sottoinsiemi e funzioni caratteristiche può quindi essere espresso dicendo che per ogni sottoggetto $\phi: U \to A$ esiste un morfismo caratteristico $\chi_U : A \to \{V,F\}$, tale che $\chi_U \cdot \phi = true \cdot\, !_U$ e il diagramma commutativo associato a quest'equazione è universale, nel senso che ogni altro diagramma $\chi_U \cdot \psi =$ true $\cdot\, !_X$, per qualsiasi $\psi: X \to A$, è tale che esiste sempre un solo $h: X \to U$ tale che $\psi = \chi_U \cdot h$ e ovviamente $!_X = !_U \cdot h$.

$$
\begin{tikzcd}
X \arrow[dr, dashed] \arrow[drr, "\psi"] \arrow[ddr, "!_X"'] & & \\
& U \arrow[r] \arrow[d, "!_U"] & A \arrow[d, "\chi_U"] \\
& 1 \arrow[r, "true"'] & \{V,F\}
\end{tikzcd}
$$

In categorie diverse da **Set**, formate per esempio da strutture algebriche variabili in modo continuo su uno spazio, il modo in cui le parti di un oggetto sono identificate non è necessariamente lo stesso che in **Set.** Così al posto di {V,F} possiamo avere un oggetto Ω con molti più valori di verità, strutturati in modo più complesso. Una categoria che ha un oggetto Ω che soddisfa alla proprietà illustrata nel diagramma precedente si dice che ha un «classificatore di sottoggetti». Un *topos elementare* è appunto una categoria cartesiana chiusa con classificatore di sottoggetti. Ci sono molti topos (il termine, essendo un acronimo, non è tenuto ad aver il plurale greco: topoi). Ogni topos può servire come un universo per fare matematica, in funzione di specifiche proprietà aggiuntive. Ci sono topos particolarmente adatti per la computabilità e altri per l'analisi. Ebbene, Lawvere ha scoperto che **Set** è contraddistinta da assiomi puramente categoriali. **Set** è infatti un topos elementare che in più soddisfa alle condizioni seguenti:

- il principio di *estensionalità*, ora nella forma: per ogni due morfismi $f,g: A \to B$, se $\forall x{:}1 \to A\ (fx = gx)$ allora $f = g$.

- la *bivalenza*, cioè $\Omega \cong \{V,F\}$
- l'assioma di *scelta*, ora riformulabile come: per ogni epimorfismo $f\colon A \to B$ esiste un $g\colon B \to A$ tale che $fg = 1_B$ (esiste una sezione di f).

Ci sono topos che non soddisfano a nessuna di queste tre proprietà aggiuntive e altri che soddisfano a qualche loro versione modificata. La teoria dei topos è una delle principali branche della teoria delle categorie ma non la esaurisce.

4. Questioni filosofiche

Le prospettive aperte dalla teoria delle categorie sono vaste e molto articolate. Ma, accanto alle difficoltà che si devono ancora superare per servirsene come strumento nelle scienze naturali (a partire dalla fisica), ci sono anche difficoltà di carattere fondazionale e filosofico, nel momento in cui la teoria si propone come alternativa alla tradizionale teoria degli insiemi, basata sul concetto di appartenenza.

Un'obiezione ricorrente è quella che rimprovera alla teoria delle categorie di dipendere da assunzioni di tipo insiemistico. Ne abbiamo visto un esempio allorché siamo arrivati a una categoria 'grande' com'è appunto la categoria di tutte le categorie. Ma si potrebbe più semplicemente notare che una categoria è definita come totalità di oggetti e morfismi, dunque come insieme.

A quest'obiezione sono state date risposte diverse, tutt'altro che convergenti. C'è chi si accontenta di una nozione intuitiva di collezione, non necessariamente catturata dalla teoria assiomatica degli insiemi, chi si preoccupa di misurare l'impegno che il riferimento alla nozione di insieme comporta, e chi sostiene che la dipendenza è solo apparente, perché dispensabile nella metateoria.

Personalmente, ritengo che la terza opzione sia quella da perseguire, adottando la linea del matematico francese Jean Benabou, cfr. [2]: molte costruzioni categoriali comportano il rimando a famiglie indiciate di oggetti, ma non è affatto detto che l'indiciazione debba essere in termini di insiemi, come oggetti di una particolare categoria. Questa linea porta al concetto di fibrazione, cruciale al pari di quello di aggiunzione nell'architettura della matematica. Purtroppo, il concetto di fibrazione richiede molti presupposti, che qui non è possibile fornire (il concetto di spazio di fibre in geometria algebrica e quello di fibrato tangente in fisica sono esempi da cui partire, ma non consentono di raggiungere il livello di generalità richiesto).

Come l'ultima questione toccata suggerisce, la filosofia della matematica che si appoggia alla teoria delle categorie non è univoca. Nondimeno, essa differisce dalle principali scuole fondazionali del Novecento: logicismo, formalismo e intuizionismo. Inoltre, non è vincolata al privilegio accordato dagli informatici alle procedure computazionali, almeno nel senso in cui la nozione di calcolabilità ha trovato sistemazione a partire dagli anni trenta (funzioni ricorsive, macchine di Turing, lambda-calcolo); e tanto meno è riducibile a una qualche teoria assiomatica degli insiemi, senza con ciò escludere minimamente l'importanza del concetto di insieme. È un modo di pensare la matematica *e* di fare matematica, che ha permesso un approfondimento di concetti come *spazio, numero, verità logica, grandezza, figura*, cfr. [3], portando a ridisegnare l'impianto delle conoscenze matematiche già acquisite e aprendo nuove strade alla ricerca in molteplici settori.

Nel quadro categoriale l'incremento sostanziale di unità, ottenuto grazie a legami espressi in forma di aggiunzioni, si accompagna al riconoscimento che le sorgenti del pensiero matematico non sono di un unico tipo. Anzi, la stessa idea dei fondamenti cambia. Dall'idea della matematica come un edificio a forma di piramide rovesciata, che poggia sul vertice (costituito da assiomi globali che abbracciano tutte le possibili totalità e poi non permettono di spiegare come mai, fra esse, proprio quelle corrispondenti al reale pensiero matematico si sono realizzate), si passa all'idea di molteplici unità modulari, combinate secondo principi universali di correlazione.

Si potrebbe anche dire che, come da un sistema di riferimento privilegiato, qual era quello identificato da Newton in uno spazio e in un tempo assoluti, si è passati allo spazio-tempo relativistico e a sistemi di riferimento collegati da principi di invarianza, così si passa da un universo omogeneo di entità matematiche a una molteplicità di universi collegati da principi trasversali, individuati appunto nella teoria delle categorie.

Il che non significa, sul piano didattico, che questi principi possano essere colti, apprezzati e impiegati sotto vuoto e senza precedenti conoscenze. Al pari dei principi della teoria degli insiemi, possono essere gradualmente introdotti di pari passo con la conoscenze via via acquisite nei vari ambiti della matematica. Ma soprattutto, a differenza quanto meno del modo in cui l'*insiemistica* è impiegata (con funzione di igiene linguistica, dispensabile operativamente quando si studiano grandezze geometriche e si fanno calcoli), l'impostazione categoriale può orientare la presentazione, da parte del docente, di ciò che è centrale in ciascun ambito, perché così facendo si facilita la comprensione di come concetti di natura diversa (come quelli della logica, dell'algebra, della geometria, dell'analisi e infine della fisica matematica) si integrano fra loro in modi nient'affatto arbitrari.

Note

[*] Il presente testo è una versione ampliata della 'lezione-incontro' tenuta a Livorno il 18 ottobre 2005, nell'ambito delle manifestazioni di *Pianeta Galileo*. Ringrazio gli allievi e i docenti dell'ITG Buontalenti per l'interesse mostrato.

BIBLIOGRAFIA

[1] Bell, J. L., *A primer of infinitesimal analysis*. Cambridge University Press, Cambridge 1998.

[2] Benabou, J., Fibered categories and the foundations of naive category theory, *Journal of symbolic logic*, 50, 1985, pp. 10-37.

[3] Lawvere, F. W., Categorie e spazio, *Lettera matematica Pristem*, 31, 1999, pp. 35-50.

[4] Lawvere, F. W., Schanuel, S., *Teoria delle categorie: un'introduzione alla matematica*, Franco Muzzio, Padova 1994.

[5] Mac Lane, S., Birkhoff, G., *Algebra*, Mursia, Milano 1975.

[6] Mangione, C. Logica e teoria delle categorie, in *Storia del pensiero filosofico e scientifico,* a cura di L. Geymonat, vol. 7, pp. 519-653, Garzanti, Milano 1976.

[7] Peruzzi, A., Il contenuto della forma logica, in *Forma e contenuto*, a cura di R. Lanfredini, pp. 211-222, LED, Milano 2002.