CAPITOLO 5

CALCOLO DELLE IDEE: LINGUAGGIO E RISCRITTURA

La matematica è un gioco che segue alcune semplici regole giocato con segni senza senso sulla carta. (*David Hilbert* 1862-1943)

1. Nuovo ruolo del linguaggio: calcolare con le parole

Fino ad ora abbiamo visto i calcolatori come strumenti per fare calcoli numerici. Se si limitassero a questo, i calcolatori moderni non avrebbero un comportamento molto diverso dagli antichi abaci o dalle macchine di Pascal. In realtà le potenzialità della moderna informatica vanno ben oltre i semplici calcoli numerici fino a giungere ad attività che simulano in qualche modo il comportamento intelligente degli esseri umani. Le ricerche scientifiche che si occupano di questo argomenti sono comprese sotto il nome di *Intelligenza artificiale.* In questo e nel prossimo capitolo tratteremo quegli aspetti dell'intelligenza artificiale che riguardano l'elaborazione dell'informazione non numerica (sistemi di riscrittura, logica matematica) e che si basano sulla manipolazione delle parole di un dato linguaggio formale. Negli altri capitoli ci occuperemo di altri aspetti dell'intelligenza artificiale che vanno sotto il nome di reti neurali, logica fuzzy, algoritmi genetici ed altro.

Come vedremo nella possibilità da parte di una macchina di riprodurre un comportamento intelligente un ruolo fondamentale viene svolto dal linguaggio. Infatti, come vedremo, nell'intelligenza umana il linguaggio gioca un ruolo centrale non solo come strumento di comunicazione ma anche come elemento costitutivo dell'intelligenza. Tale ruolo ha origini lontane: ne evidenzio alcune tappe.

La notazione posizionale al posto dell'abaco. Ad esempio se si considera la notazione posizionale per la rappresentazione dei numeri ci si accorge subito che essa fa assumere al linguaggio una valenza completamente nuova. Infatti è da questo momento che il linguaggio matematico:
<u>da semplice strumento per potere comunicare le idee di natura matematica, diviene oggetto di manipolazione e calcolo.</u>
Se scrivo 237 non uso solo un simbolo per denotare un numero. Piuttosto fornisco informazioni su come tale numero può essere scomposto in somma di potenze di 10. Quando poi effettuo una operazione, ad esempio 237 + 45, in effetti non lavoro sui numeri 237 e 45 ma su una loro rappresentazione (cioè su due parole in un opportuno linguaggio). Il risultato non è solo un numero, ma una rappresentazione, cioè la parola 282 in cui si fornisce anche informazione su come tale numero possa essere scomposto in somma di potenze di dieci. Allora i calcoli con i numeri in notazione posizionale sono in effetti elaborazione di parole secondo certe regole. E' anche interessante osservare che la scrittura di un numero equivale a collocare opportunamente le palline sulle varie aste di un abaco e che l'esecuzione dell'addizione, compreso il riporto, simula lo spostamento delle palline con cui in un abaco viene eseguita l'addizione.

Dall'algebra retorica al calcolo letterale. Un altro momento importante è la nascita dei metodi algebrici agli inizi del 1600, metodi caratterizzati dall'uso sistematico del calcolo letterale.

Si osservi che ancora nel 1500 lo studio delle equazioni algebriche era di tipo "retorico". Questo significava che la descrizione di una equazione da risolvere veniva data nel linguaggio naturale, senza cioè l'uso di lettere per indicare l'incognita e lettere per indicare i coefficienti dei polinomi. Non era quindi possibile procedere ad una trattazione generale delle equazioni algebriche. Mancavano perfino simboli a noi familiari come +, -, = , x^2 ... Un esempio di algebra retorica è dato dal famoso sonetto in cui Tartaglia enuncia la regola risolutiva della equazione di terzo grado. Il sonetto comincia al modo seguente:

Quando che il cubo con le cose appresso
Se agguaglia à qualche numero discreto
...

Versi questi che si "traducono" attualmente nell'equazione:

x^3	+	ax	=	n
cubo	con appresso	cose	agguaglia	a numero

Un altro esempio per renderci conto delle difficoltà che si dovevano affrontare per la mancanza di un linguaggio simbolico adeguato, è fornito dal seguente brano di un matematico italiano del tempo, Luca Pacioli (Borgo S. Sepolcro 1445 ca. -1514 ca.).

Exēplum al.3°.
cõposto.Trouame.1.numero.che multiplicato per.5.facia quāto el suo q̄drato giõto cõ.4.
Poi ch'l sia.1.co.el suo quadrato e ne.1.ce.giontoci.4.sira eq̄le a.5.via.1.co.cioe.1.ce.p̄.4.se
agualiano a.5.co. Smeçça le cose.Multiplica in se.Lauane el numero.Restara.2 ¼.e la.℞.
2 ¼.p̄.2 ½.per lo dimeçamento de le cose valse la cosa.E fo el domandato numero: cioe.4.

Trascriviamo il brano letteralmente:

> *Exemplum al 3º composto.*
> *Trovame 1 numero che multiplicato per 5 facia quanto el suo quadrato gionto con 4.*
> *Poni ch'l sia 1 co.; el suo quadrato e ne 1 ce. giontoci 4 sirà equale a 5 via 1 co.; cioè 1 ce. p. 4 se aguagliano a 5 co.*
> *Smezza le cose. Multiplica in sé. Lavane el numero. Restarà 2 1/4 e la R. 2 1/4 p. 2 1/2, per lo dimezzamento delle cose, valse la cosa. E fo el domandato numero: cioè 4.*

Nel primo rigo Pacioli inizia con l'enunciare il problema:

Trovare un numero che moltiplicato per 5 sia uguale al suo quadrato più 4.

Si riferisce all'equazione $5x = x^2+4$ che però sente il bisogno di riscrivere chiamando l'incognita "la cosa" e abbreviata in "1 co."ed il quadrato dell'incognita (per noi x^2) con un altro nome, "censo", e abbreviato in "1 ce."

Poi, quasi rendendosi conto della difficoltà ad interpretare il problema posto, ripete ancora una volta la condizione da imporre ad x dicendo:
1 ce. p. 4 è uguale a 5 co.
cioè "un censo più 4 è uguale a 5 via (per) una cosa".

Una volta giunti all'equazione seguono poi le istruzioni per risolverla. Per capirle ricordiamo che la formula risolutiva dell'equazione di secondo grado $x^2 - bx + c = 0$ ci dice che una soluzione (l'altra viene qui trascurata) è

$$\frac{5+\sqrt{25-16}}{2} = \frac{5}{2} + \sqrt{\left(\frac{5}{2}\right)^2 - 4} = \left(2+\frac{1}{2}\right) + \sqrt{\left(2+\frac{1}{2}\right)^2 - 4}$$

E questo è proprio ciò che dice Pacioli quando dice

> *Smezza le cose. Multiplica in sé. Lavane el numero. Restarà 2 1/4 e la R. 2 1/4 p. 2 1/2, ... E fo el domandato numero: cioè 4.*

Infatti, se proviamo a tradurre quanto detto abbiamo:

"Smezza le cose", cioè dividi per due il coefficiente dell'incognita, 5. Viene 2+1/2.
"Multiplica per sé". Viene $(2+ 1/2)^2$.
"Levane el numero", cioè sottrai 4
"Restarà 2 1/4. si ottiene $(2+ 1/2)^2 – 4 = 2+1/4$.
"e la R. 2 ¼ p. 2 1/2" – e la radice quadrata (indicata con una R tagliata) di 2 + 1/4 più 2 + 1/2, cioè $\sqrt{(2+1/4)} + 2 + 1/2$

"*valse la cosa*" cioè si ottiene il numero $\sqrt{2+\frac{1}{4}}+2+\frac{1}{2}$.

"E fo el domandato numero: cioè 4" che risulta uguale al numero cercato, cioè 4.

Si vede in questo scritto di Pacioli tutta la difficoltà che si incontra a giungere ad una trattazione delle equazioni algebriche quando non si usi un linguaggio algebrico appropriato (osserveremo invece come per Aristotele l'uso delle variabili per denotare proposizioni appaia completamente naturale).

Un forte passo avanti verso il moderno simbolismo fu fatto da Viète (Francia, 1540-1603) il quale usò lettere per rappresentare le incognite e lettere per rappresentare i coefficienti dei polinomi. Ciò ebbe due effetti positivi.

1. Liberò l'algebra dalla considerazione di equazioni particolari (cioè con particolari coefficienti) e quindi diede la possibilità sia di procedere in modo generale. Si osservi che prima del calcolo simbolico ogni equazione veniva considerata un caso a parte.
2. Creò un vero e proprio calcolo delle equazioni. Tale calcolo, che attualmente viene chiamato anche "sistema di riscrittura", permette di passare da una equazione ad un'altra più semplice fino a giungere ad una equazione del tipo $x = m$ in cui l'incognita compare solo alla sinistra dell'equazione ed in maniera esplicita.

2. Con Cartesio non solo l'algebra ma anche la geometria diventa calcolo.

Quando appare la *Geometria* di Cartesio nel 1637 il calcolo simbolico dell'algebra ha raggiunto la sua piena maturità. In tale opera Cartesio inizia un sistematico processo di algebrizzazione della geometria che terminerà nella attuale geometria analitica. La geometria "sintetica" di Euclide, in cui le dimostrazioni giocano un ruolo centrale, viene tradotta da Cartesio in uno studio delle equazioni algebriche in cui tutti i problemi si riducono al calcolo delle radici di sistemi di equazioni algebriche. In un certo senso si passa dalle dimostrazioni con figure tipiche della geometria euclidea ai calcoli tipici dell'aritmetica: da tale punto di vista non è sbagliato considerare il metodo di Cartesio come una forma di automatizzazione dei ragionamenti. Da notare che Cartesio non amava molto la logica che identificava con la teoria dei sillogismi di Aristotele. Infatti vedeva tale logica come qualcosa di sterile. Ad esempio dice:

> *Quanto alla logica, i suoi sillogismi e la maggioranza dei suoi altri precetti sono utili piuttosto per la comunicazione di quanto già conosciamo . . . piuttosto che per l'investigazione di ciò che ci è sconosciuto.*

Tuttavia di fatto le sue idee costituiscono un passo importante verso "l' algebrizzazione e meccanizzazione del ragionamento" e quindi, di fatto, il primo passo significativo verso la logica matematica. Scopo dichiarato di Cartesio è la ricerca di un metodo generale in contrasto con il modo frammentario con cui procedevano i greci antichi quando si trattava di trovare una dimostrazione o di risolvere un problema:

> *Quanto poi all'analisi degli antichi e all'algebra dei moderni . . . , la prima è sempre siffattamente legata alla considerazione delle figure, che essa non può esercitare l'intelligenza senza affaticare di molto l'immaginazione; e, nella seconda, ci si è talmente assoggettati a certe regole e a certe cifre, che se ne è fatta un'arte confusa ed oscura, la quale tiene imbarazzato lo spirito, invece di (essere) una scienza che lo coltivi.*

Se infatti è certamente un merito dei greci il fatto che ogni dimostrazione sia rigorosamente controllabile nei sui singoli passaggi, niente viene detto da essi circa il "metodo", cioè l'algoritmo, che si dovrebbe seguire per poter trovare nuovi teoremi e dimostrazioni. Pertanto restiamo disarmati di fronte ad ogni problema nuovo che si presenta e dobbiamo ogni volta inventarci una soluzione.

Il metodo proposto da Cartesio per la geometria consisteva

- nell'indicare con lettere i dati e le incognite di un problema geometrico (introduzione di un inguaggio formale)
- di tradurre le informazioni disponibili in equazioni (rappresentazione dell'informazione: un analogo del sistema di assiomi)
- nel semplificare, tramite calcoli algebrici, tali equazioni quanto più possibile fino a giungere a formule risolutive (processo deduttivo che in termini moderni corrisponde ai *rewriting systems*)
- nell'interpretare le formule risolutive in termini di calcolo dei segmenti.

> *Così, volendo risolvere qualsiasi problema, si deve innanzi tutto considerarlo come risolto, e si devono dare di nomi a tutte le linee che sembrano necessarie per la sua costruzione, sia quelle ignote, sia alle altre. Poi, senza fare alcuna differenza tra queste linee, note ed ignote, bisogna affrontare le difficoltà secondo l'ordine che mostra nella maniera più naturale in che modo tali linee siano in rapporto tra loro, fino a che non si sia trovato modo di esprimere una medesima quantità in due maniere diverse: ciò si chiama un'equazione (in una sola incognita) poiché i termini di una di queste due espressioni sono uguali a quelli dell'altra.*

Da notare che *La Geometria* non si presenta come un libro a se stante che abbia come unico scopo la ricerca di risultati scientifici su di un particolare argomento. La Geometria è una delle appendici del famoso *Discorso sul Metodo* del 1637 il cui titolo completo è

"Discorso sul metodo per ben condurre la propria ragione e cercare la verità nelle scienze".

Tale titolo si potrebbe tradurre in

"Discorso sulla logica"

e quindi non è azzardato dire che la geometria di Cartesio sia un capitolo della logica purché alla parola logica si dia un significato più ampio di quello dato da Aristotele[1].

2. Cartesio e Leibniz: inutile affaticarsi a discutere, calcoliamo

Pertanto quello che Cartesio proponeva era una riduzione della geometria ai suoi elementi più semplici, i segmenti, e non una riduzione della geometria a manipolazione di numeri come avviene attualmente.

> *Tutti i problemi della geometria si possono facilmente ridurre a tali termini che in seguito per costruirli basta conoscere la lunghezza di alcune rette.*

Come abbiamo visto tali elementi semplici, cioè i segmenti, si possono manipolare con operazioni simili a quelle dell'aritmetica. Pertanto è corretto dire che con Cartesio si ha una algebrizzazione della geometria purché con tale espressione non si intenda riduzione alla teoria dei numeri reali ma, in accordo il punto di vista più moderno, come introduzione di un punto di vista algoritmico-manipolatorio che pone la nozione di operazione alla base di tutto. Scopo dichiarato di Cartesio è la ricerca di un metodo generale in contrasto con il modo frammentario con cui procedevano i greci antichi quando si trattava di trovare una dimostrazione o di risolvere un problema. Infatti è da notare che *La Geometria* non si presenta come un libro a se stante che abbia come unico scopo la ricerca di risultati scientifici su di un particolare argomento. La Geometria è una delle appendici del famoso *Discorso sul Metodo* del 1637 il cui titolo completo è *"Discorso sul metodo per ben condurre la propria ragione e cercare la verità nelle scienze"* E' quindi uno strumento per illustrare il suo metodo filosofico generale. Ecco cosa dice Cartesio:

---

[1] Resta il fatto che l'accusa di sterilità fatta da Cartesio alla logica pone in evidenza un limite anche della moderna logica matematica quando, coerentemente con le idee di D. Hilbert, viene vista come strumento di sistemazione rigorosa delle conoscenze e non strumento di scoperta. Tuttavia le attuali ricerche svincolano quasi completamente la logica dal suo originale legame con il problema dei fondamenti della matematica. Si veda ad esempio la teoria dei modelli, la teoria della calcolabilità, la logica per l'intelligenza artificiale.

> *Quanto poi all'analisi degli antichi e all'algebra dei moderni . . . , la prima è sempre siffattamente legata alla considerazione delle figure, che essa non può esercitare l'intelligenza senza affaticare di molto l'immaginazione; e, nella seconda, ci si è talmente assoggettati a certe regole e a certe cifre, che se ne è fatta un'arte confusa ed oscura, la quale tiene imbarazzato lo spirito, invece di (essere) una scienza che lo coltivi.*

Se infatti è certamente un merito dei greci il fatto che ogni dimostrazione sia rigorosamente controllabile nei sui singoli passaggi, niente viene detto da essi circa il "metodo", cioè l'algoritmo, che si dovrebbe seguire per poter trovare nuovi teoremi e dimostrazioni. Pertanto restiamo disarmati di fronte ad ogni problema nuovo che si presenta e dobbiamo ogni volta procedere per tentativi.
Il metodo proposto da Cartesio per la geometria consisteva

- nell'indicare con lettere i dati e le incognite di un problema geometrico
- di tradurre le informazioni disponibili in equazioni
- nel semplificare, tramite calcoli algebrici, le equazioni quanto più possibile fino a giungere a formule risolutive
- nell'interpretare le formule risolutive in termini di calcolo dei segmenti.

Ecco che cosa dice Cartesio nella sua *Geometria*

> *Così, volendo risolvere qualsiasi problema, si deve innanzi tutto considerarlo come risolto, e si devono dare di nomi a tutte le linee che sembrano necessarie per la sua costruzione, sia quelle ignote, sia alle altre. Poi, senza fare alcuna differenza tra queste linee, note ed ignote, bisogna affrontare le difficoltà secondo l'ordine che mostra nella maniera più naturale in che modo tali linee siano in rapporto tra loro, fino a che non si sia trovato modo di esprimere una medesima quantità in due maniere diverse: ciò si chiama un'equazione (in una sola incognita) poiché i termini di una di queste due espressioni sono uguali a quelli dell'altra.*

Il potere dell'algebra di meccanizzare gli argomenti geometrici ben presente in Cartesio suggestionò poi anche il grande filosofo e matematico G. W. Leibniz. Entrambi prefiguravano una scienza più ampia dell'algebra dei numeri. Tale scienza avrebbe operato con lo stesso metodo dell'algebra astratta ma sarebbe stata applicabile ai ragionamenti in tutti i campi della conoscenza razionale e non solo a quelli della geometria. Vediamo ad esempio che cosa dice Leibniz:

> *Ogni ragionamento umano si compie per mezzo di certi segni o caratteri ... le lingue ordinarie, sebbene servano al ragionamento, tuttavia sono soggette ad innumerevoli equivoci, né possono essere impiegate per il calcolo ... Questo mirabilissimo vantaggio finora lo danno soltanto i segni impiegati dagli aritmetici e dagli algebristi, nei quali ogni ragionamento consiste nell'uso di caratteri, e ogni errore mentale è lo stesso che un errore di calcolo ...*
> *... mi apparve subito chiaro che tutti i pensieri umani possono risolversi del tutto in pochi pensieri da considerarsi come primitivi. Se poi si assegnano a questi ultimi dei caratteri, di qui si possono formare i caratteri delle nozioni derivate ... Una volta fatto questo, chiunque si servisse dei caratteri così descritti nel ragionare nello scrivere, o non commetterebbe mai errori, oppure li riconoscerebbe sempre da sé, siano suoi o degli altri, mediante esami facilissimi.*
> *... una volta fatto ciò, quando sorgeranno delle controversie, non ci sarà maggior bisogno di discussione tra due filosofi di quanto ce ne sia tra due persone che effettuano calcoli. Sarà sufficiente, infatti, che essi prendano la penna in mano, si siedano ad un tavolino, e si di dicano reciprocamente "calcoliamo".*

Leibniz pertanto suggerisce che il ragionamento si può ridurre ad un calcolo dei caratteri, cioè ad un calcolo in cui gli oggetti non sono più numeri ma parole, per la precisione parole di un linguaggio appositamente costruito. Detto questo, visto che ragionare è manipolare in modo effettivo parole di un linguaggio, il primo passo per la costruzione di macchine che ragionano è quello di insegnare alle macchine in modo preciso che cosa è un linguaggio e poi di definire una nozione di "calcolo" su di un dato linguaggio.
A differenza delle idee relative al calcolo differenziale, la proposta di Leibniz di un calcolo logico non ebbe molti sviluppi. Si dovranno aspettare gli inizi del novecento per la moderna formulazione, da parte del grande matematico D. Hilbert, della logica matematica.

3. Parole linguaggi e grammatiche: macchine che parlano correttamente

Se vogliamo precisare che cosa si debba intendere per calcolo dei concetti e per logica matematica, dobbiamo dare una definizione rigorosa di che cosa si debba intendere per *"parola"* e per *"linguaggio"*. Inoltre dobbiamo estendere ai linguaggi le nozioni fondamentali di computabilità, di effettiva enumerabilità e di decidibilità che fino ad ora abbiamo dato solo facendo riferimento all'insieme dei numeri naturali. Questo perché abbiamo utilizzato la nozione di funzione parziale ricorsiva e la nozione di macchina a registri ed entrambe queste nozioni si riferiscono ai numeri interi. D'altra parte abbiamo già dato definizioni di macchine che lavorano con inputs ed outputs non numerici. Ad esempio gli automi finiti e le macchine di Turing lavorano con successioni di caratteri di un dato alfabeto, caratteri non necessariamente numerici. In particolare un automa trasforma sequenze in sequenze e non solo elementi in elementi come nel caso delle funzioni. Per tale motivo spesso gli automi finiti vengono detti anche "macchine sequenziali". La nozione di "parola" su un dato alfabeto permette di esprimere meglio tale azione di un automa. Cominciamo ad esaminare la nozione usuale di parola. E' chiaro che una parola come "neve" o come "cane" è costituita da lettere di un alfabeto A ma è anche chiaro che tali parole non sono un semplice insieme di lettere. In effetti la nozione di insieme non è adeguata in quanto non tiene conto dell'ordine in cui sono state scritte le lettere e non tiene conto delle eventuali ripetizioni. Se volessimo rappresentare la parola "cane" con $\{c,a,n,e\}$ non sapremmo distinguerla dalla parola "cena" in quanto $\{c,a,n,e\}$ e $\{c,e,n,a\}$ sono due insiemi uguali. In definitiva si deve poter dire quale è la prima lettera della parola, quale la seconda e così via. La teoria degli insiemi fornisce la nozione di *n-pla* che è sicuramente adeguata in quanto, appunto, in una n-pla è possibile distinguere il primo elemento, il secondo e così via. In definitiva dovendo definire in maniera rigorosa la nozione di parola dovremmo procedere al modo seguente:

Definizione 1. Sia A un insieme finito che chiamiamo *alfabeto*, allora una *parola* di lunghezza n su A è una n-pla di elementi di A. Indichiamo con A^+ l'insieme delle parole sull'alfabeto A, cioè poniamo:

$$A^+ = \cup_{n\in N} A^n.$$

Invece di usare la notazione insiemistica $(a_1,...,a_n)$ per indicare una parola, scriveremo semplicemente $a_1...a_n$. Ad esempio se A è l'alfabeto $\{a,c,d\}$ allora scriveremo ad, aac, caa per denotare le parole (a,d), (a,a,c) e (c,a,a) rispettivamente.

Il concetto di linguaggio nasce quando sia stato indicato qualche modo per distinguere le parole "sensate" da quelle "non sensate" (quelle scritte "secondo le regole" da quelle "scritte male"); quindi, se si assume il punto di vista *estensionale* tipico della teoria degli insiemi, quando si sia definito un sottoinsieme di A^*.

Definizione 2. Chiamiamo *linguaggio formale* sull'alfabeto A ogni sottoinsieme $\mathcal{L}$ di A^+.

Un linguaggio può essere dato, nel caso sia finito, mediante un elenco delle parole che lo compongono. Un esempio è fornito dall'insieme delle parole che sono indicate nel vocabolario italiano. Comunque in genere un linguaggio viene dato tramite un insieme di regole capaci di produrre i suoi elementi e pertanto può essere anche infinito. Ad esempio, è possibile costruire gli elementi di un linguaggio "dal basso" indicando esplicitamente alcune parole-base di $\mathcal{L}$ ed alcuni procedimenti per costruire nuovi elementi di $\mathcal{L}$ a partire da dati elementi di $\mathcal{L}$.

Proposizione 3. L'insieme A^+ è numerabile, pertanto ogni linguaggio è finito o numerabile.

Dim. Supponiamo che l'alfabeto A sia costituito da p elementi. Allora l'insieme $P(n)$ delle parole di lunghezza n è finito in quanto contiene p^n elementi. Poiché $A^* = \cup_{n\in N}P(n)$, l'insieme A^* risulta numerabile (si ricordi che l'unione di una successione di insiemi finiti o numerabili è un insieme finito o numerabile). In particolare ogni linguaggio su A, in quanto sottoinsieme di A^*, risulta finito o numerabile.

Esempio. Con riferimento all'alfabeto $A = \{x, y, +, \cdot, (,)\}$, potremmo avere le seguenti regole che definiscono il linguaggio dei polinomi nelle variabili x ed y;

i) x ed y sono elementi di $\mathcal{L}$

ii) se p_1 e p_2 sono elementi di $\mathcal{L}$ allora anche $(p_1)\cdot(p_2)$, e $(p_1)+(p_2)$ sono elementi di $\mathcal{L}$.

Il linguaggio dei polinomi viene definito come il più piccolo sottoinsieme di A^+ contenente x ed y e chiuso rispetto la regola ii).

Si noti che il termine "parola" deve essere inteso in un senso più ampio dell'usuale tanto da comprendere anche quelle che noi chiamiamo usualmente frasi, cioè sequenze di parole. Infatti può capitare che l'alfabeto *A* sia a sua volta costituito da parole su di un altro alfabeto *B*. In tale caso una parola nell'alfabeto *A* viene ad essere una sequenza di parole nell'alfabeto *B* per evitare confusioni tra i due alfabeti indicheremo le parole nell'alfabeto *A* interponendo uno spazio vuoto tra elementi successivi di *B*.

Un linguaggio *L* può essere dato, nel caso sia finito, mediante un elenco delle parole che lo compongono, ma in genere un linguaggio *L* viene dato tramite un insieme di regole capaci di produrre i suoi elementi e pertanto può essere anche infinito. Ad esempio, è possibile costruire gli elementi di un linguaggio "dal basso" indicando esplicitamente alcune parole-base di *L* ed alcuni procedimenti per costruire nuovi elementi di *L* a partire da dati elementi di *L*.

Uno dei modi principali per generare un linguaggio è tramite le *grammatiche*. Da notare che la nozione di grammatica non è nata nell'ambito informatico. Essa è stata proposta per la prima volta dal linguista americano Noam Chomsky nel 1957. La questione affrontata da Chomsky è quella di capire l'origine della straordinaria capacità dei bambini di scoprire le "regole" che stanno alla base del linguaggio. E questo a partire da un numero limitatissimo di emissioni verbali dei genitori e delle persone che li circondano. Questa capacità è tanto più sorprendente quanto si consideri che da un numero finito di casi i bambini apprendono il modo di produrre un numero potenzialmente infinito di frasi. La nozione di grammatica fornisce in un certo senso una risposta a questioni di tale tipo, essendo, come vedremo, una grammatica un "oggetto finito" capace di produrre un numero potenzialmente infinito di frasi. Se si accetta questo punto di vista, non è strano che un bambino, in base al numero limitato delle esperienze linguistiche assorbite dall'ambiente, sia in grado di apprendere le regole della grammatica della lingua dei genitori.

Definizione 4. Una *grammatica* è una quadrupla $G = (V,A,P,s)$ con:
- V (insieme delle *variabili ausiliari*) ed A (insieme dei *simboli terminali*) insiemi finiti disgiunti;
- s elemento prefissato di V detto *start-simbolo*;
- P insieme finito di espressioni del tipo $\alpha\to\beta$ (dette *produzioni*) con α e β parole nell'alfabeto $C = V\cup A$.

Ogni produzione $\alpha\to\beta$ va intesa come possibilità di sostituire in una parola alcune occorrenze di α con β. Se la parola y si ottiene dalla parola x tramite una tale sostituzione allora si dice che y deriva direttamente da x. Più precisamente, se $\alpha\to\beta$ è un elemento di P ed δ e γ sono elementi di C^+ allora diremo che $\gamma\beta\delta$ *deriva direttamente* da $\gamma\alpha\delta$ che $\beta\delta$ *deriva direttamente* da $\alpha\delta$ e che $\gamma\beta$ *deriva direttamente* da $\gamma\alpha$.

Definizione 5. Chiameremo *derivazione* una catena finita $x_1,...,x_n$ di elementi di C^+ tali che x_{i+1} deriva direttamente da x_i per $i =1,...,n-1$; in tale caso, se $\alpha = x_1$ e $\beta = x_2$, scriveremo $\alpha \mid\!\to \beta$.

In altri termini β deriva da α nella grammatica G se è possibile ottenere β da α mediante una ripetuta sostituzione di sottoparole in accordo con quanto consentito dalle regole di produzione elencate in P.

Definizione 6. Diremo che un linguaggio L nell'alfabeto A è *producibile* dalla grammatica G se gli elementi di L si possono caratterizzare come le parole nell'alfabeto terminale A che si possono derivare dallo start simbolo s cioè se $L = \{x\in A^+ : s \mid\!\to x\}$. Chiamiamo *complessità* di un elemento x di L la minima lunghezza delle derivazioni che consentono di ottenere x.

Per avere una notazione più breve, se esistono più produzioni $\alpha \to \beta_1$ $\alpha \to \beta_2$, . . . $\alpha \to \beta_n$ con la stessa premessa α allora indicheremo con $\alpha\to\beta_1 /\beta_2/\ldots/\beta_n$ tale insieme di produzioni. Nel seguente esempio gli alfabeti V ed A sono a loro volta insiemi di parole in un altro alfabeto.

Esempio. Vediamo come sia possibile costruire, tramite una opportuna grammatica, una piccolissima fetta della lingua italiana. Poniamo:
V = {*asserzione*, *articolo*, soggetto, *predicato*}
A={*il*, *un*, *cane*, *gatto*, *abbaia*, *corre*}
e supponiamo che le produzioni siano:

asserzione → *articolo soggetto predicato*;
articolo → *il* / *un* ;
soggetto → *cane* / *gatto* ;
predicato → *corre* / *abbaia* .

Assunto come start-simbolo la variabile "*asserzione*", il linguaggio della grammatica definita in questo modo consentirà di produrre frasi del tipo
"il cane abbaia", "il gatto abbaia", "il cane corre", "il gatto corre" .
Ad esempio, abbiamo la derivazione
asserzione, articolo cane predicato, il cane predicato, il cane abbaia
che si ottiene utilizzando le produzioni
asserzione → *articolo soggetto predicato, soggetto* → *cane, articolo*→ *il, predicato* → *abbaia*.
Non deve stupire il fatto che sia producibile una frase del tipo *"il gatto abbaia"* in quanto una grammatica non garantisce che le frasi costruite siano vere ma solo che siano scritte correttamente. Aumentando il numero di possibili soggetti (compatibili con l'articolo "*il*") e predicati (in terza persona singolare) è possibile ottenere un maggior numero di frasi della lingua italiana scritte correttamente. La costruzione di una grammatica capace di produrre (e controllare) l'intera lingua italiana è ovviamente una cosa complicatissima e non ancora completamente realizzata. Una tale realizzazione permetterebbe, ad esempio, di programmare un sistema che segnali automaticamente gli errori di grammatica di chi scrive (attualmente vengono segnalati solo gli errori di vocabolario). Tale realizzazione sarebbe inoltre necessaria per progettare un buon traduttore dall'italiano all'inglese e viceversa. D'altre parte non è affatto sicuro che la nozione che abbiamo proposto di grammatica possa consentire una tale impresa. Spesso gli aspetti semantici e quelli sintattici si intrecciano nel senso che la struttura grammaticale di una frase dipende non solo dalla sua struttura formale ma anche dal significato delle parole. Ad esempio consideriamo le frasi
Maria mangiò la pizza perché era croccante
Maria mangiò la pizza perché era affamata.
Tali frasi hanno esattamente la stessa struttura formale ma nel primo caso l'aggettivo "croccante" si riferisce alla pizza, nel secondo caso l'aggettivo "affamata" si riferisce a Maria. Il fatto è che l'opinione generale che abbiamo del mondo tende ad escludere che una pizza possa essere affamata o che Maria possa essere croccante.

Nota 1. L'esempio precedente fornisce un tipo di grammatica capace di produrre solo un numero finito di frasi. Ciò è dovuto al fatto che non ci sono "produzioni che chiamano se stesse". Possiamo modificare tale esempio aggiungendo una produzione del tipo

discorso → *asserzione e discorso* ; *discorso* → *asserzione*

dopo aver aggiunto "*discorso*" a *V* e la lettera "*e*" ad *A*. In tale caso si deve considerare come start-simbolo "*discorso*". E' evidente che, per la seconda di tali produzioni, sarà possibile produrre tutte le frasi della grammatica precedente ma che per la prima di tali produzioni sarà possibile scrivere frasi di lunghezza non determinata. Ecco un esempio di derivazione:

discorso
asserzione e discorso
asserzione e asserzione e discorso
asserzione e asserzione e asserzione
. . .
il gatto abbaia e il cane corre e il gatto abbaia.

Quando una regola richiama se stessa si dice anche che è *ricorsiva.*

Esempio. Sia G una grammatica in cui $A = \{a,b\}$, $V = \{v\}$, $P = \{v \to ava, v \to b\}$ e v è lo start-simbol. Allora una derivazione è costituita dalla successione
v, ava, aavaa, aabaa.
Pertanto la parola *aabaa* appartiene al linguaggio generato da G.

Esercizio. Dire quali delle parole *abb*, *ava*, *aaabaaa*, *aba*, appartiene al linguaggio generato dalla grammatica G ora definita e trovarne la relativa complessità. Scrivere una parola del relativo linguaggio di lunghezza maggiore di 5.
Esercizio. Trovare una grammatica capace di generare l'insieme delle parole del tipo $(ab)^n$.

Nota. Quando parliamo di un linguaggio formale siamo costretti ad usare a nostra volta un linguaggio informale (ad esempio la lingua italiana utilizzata in questi appunti). Per evitare confusioni chiameremo *linguaggio oggetto* il linguaggio formale di cui vogliamo parlare e *metalinguaggio* il linguaggio utilizzato per parlarne. Un fenomeno simile avviene quando studiamo la lingua inglese (linguaggio oggetto) utilizzando la lingua italiana (metalinguaggio). Un insegnante di italiano usa un metalinguaggio (la lingua italiana) per studiare una lingua (ancora la lingua italiana). In questo caso linguaggio e metalinguaggio coincidono ed a volte si usano le virgolette per riferirsi al linguaggio oggetto. Ad esempio si scriverà la parola "italiano" è scritta in italiano.

4. Computabilità, decidibilità e semidecidibilità di insiemi non numerici
Proviamo ora a ricollegarci alle nozioni di computabilità, decidibilità e semidecidibilità studiate nei capitoli precedenti. Per potere fare questo è utile introdurre la nozione di codifica di un linguaggio. Intuitivamente codificare un linguaggio significa assegnare un numero intero ad ogni parola del linguaggio in modo che a parole diverse corrispondano numeri diversi ed in modo che "codifica" e "decodifica" siano processi effettivi.

Definizione 1. Sia A un alfabeto finito. Chiamiamo *codifica* di A^+ la corrispondenza cod: $A^+ \rightarrow N$ ottenuta al modo seguente:
- detto p il numero di elementi di A, indichiamo con q un intero tale che $p \leq 2^q$
- scegliamo una funzione iniettiva $f : A \rightarrow \{0,1\}^q$
- associamo ad ogni parola $a_1,a_2...a_n$ il numero $1f(a_1)f(a_2)...f(a_n)$ scritto in base due.

Indico con $dec : N \rightarrow A^+$ la funzione, non ovunque definita, inversa di cod.

Ad esempio sia A il linguaggio $\{a,b,c\}$, allora possiamo considerare la funzione $f : A^+ \rightarrow \{0,1\}^2$ ottenuta, ad esempio, ponendo $f(a) = 01$, $f(b) = 00$, $f(c) = 11$. Allora una parola come *bacca* sarà codificata dal numero $cod(bacca) = 100011101$. Viceversa se ho un numero 111110001 per decodificarlo tolgo il primo uno e rimane 11110001, divido in blocchi di lunghezza due ed ottengo 11 11 00 01, sostituisco le coppie di numeri con le lettere corrispondenti ed ottengo $dec(111110001) = ccba$. Naturalmente non tutti i numeri sono codice di qualche parola e quindi la decodifica non è ovunque definita. Ad esempio 0010, non iniziando con 1, non è codice di nessuna parola. Il numero 110, non essendo 10 corrispondente di nessuna lettera, non è codice di nessuna parola. Da osservare che l'esistenza della funzione iniettiva f è conseguenza del fatto che il numero di elementi di $\{0,1\}^q$ è 2^q, e quindi la cardinalità di A è minore della cardinalità di $\{0,1\}^q$. L'avere aggiunto 1 alla sinistra della parola $f(a_1)f(a_2)...f(a_n)$ ha lo scopo di evitare il fatto che stringhe del tipo 1, 01, 001, ... vengano interpretate dallo stesso numero. Che la corrispondenza ottenuta in tale modo sia iniettiva è ovvio.

Si noti che esistono molti modi di procedere per codificare A^+ e che quello indicato nella dimostrazione è solo uno dei possibili.
Le nozioni fondamentali di computabilità, di effettiva enumerabilità e di decidibilità che abbiamo dato nei capitoli precedenti riferendoci all'insieme dei numeri naturali possono essere estese facilmente ad insiemi qualsiasi purché "codificabili" ed, in particolare, ai linguaggi. D'altra parte è noto che i calcolatori (apparentemente) fanno ben altre cose che trattare numeri interi. Ad esempio un word-processor come Microsoft Word riceve in input dalla tastiera sequenze di lettere e restituisce in output una successione di pagine che appaiono sullo schermo o vengono stampate. Tuttavia di fatto i computer attuali lavorano con numeri interi (rappresentati in base due). Spetta alle periferiche di ingresso (ad esempio la tastiera) trasformare gli inputs non numerici in inputs numerici. Spetta poi alle periferiche di uscita (ad esempio lo schermo) trasformare gli outputs numerici in outputs non numerici. La nozione di codifica e decodifica permette di estendere tutto quanto detto fino ad ora con riferimento all'insieme dei numeri naturali ad insiemi qualunque purché tali insiemi siano codificabili.

Definizione 3. Siano A un alfabeto finito, allora una funzione f: $A^+ \rightarrow A^+$ si dice *computabile* se esiste una funzione computabile $g : N \to N$ tale che per ogni $x \in A^+$,

$$f(x) = decod(g(cod(x))).$$

Possiamo anche dire che f è computabile se esiste una funzione computabile $g : N \to N$ tale che il seguente diagramma *commuta*:

$$\begin{array}{ccc} A^+ & \xrightarrow{cod} & N \\ f\downarrow & & g\downarrow \\ A^+ & \xleftarrow{decod} & N \end{array}$$

In altre parole per calcolare $f(x)$
- devo prima trasformare x nel numero $cod(x)$,
- poi devo effettuare opportuni calcoli tramite g ed ottenere il risultato numerico $g(cod(x))$,
- infine devo decodificare il risultato $g(cod(x))$ cioè devo interpretarlo come elemento di A^+.

Possiamo supporre che la funzione g sia calcolata da una macchina a registri opportunamente programmata, che la funzione *cod* corrisponda alla periferica di ingresso e la funzione *dec* alla periferica d'uscita.

Questa definizione corrisponde al modo di procedere di un calcolatore. Consideriamo ad esempio un programma che "rovescia" le parole, accettando ad esempio la parola "*cane*" come input e restituendo la parola "*enac*". Allora il calcolatore quando viene fornito tramite la tastiera l'input "*cane*" lo trasforma in una sequenza di 0 ed 1, cioè in un numero intero. Dopo avere lavorato su questo numero intero (tramite la funzione g) giunge al valore $g(cod(cane))$ che è ancora una sequenza di 0 ed 1, cioè un numero intero scritto in base 2. Infine deve decodificare tale intero restituendo sullo schermo la parola $enac = decod(g(cod(cane)))$.

Definizione 4. Un sottoinsieme X di A^+ è detto *decidibile* (*effettivamente enumerabile*) se l'insieme dei suoi numeri di codice $cod(X)$ è decidibile (effettivamente enumerabile).

Il fatto che X sia decidibile significa allora che esiste un programma capace di calcolare la funzione caratteristica f di $cod(X)$. Questa definizione rappresenta esattamente ciò che viene fatto da un calcolatore. Per fare un esempio, consideriamo il problema di controllare se una parola e "palindroma", cioè se si possa leggere sia da destra verso sinistra che da sinistra verso destra. Le parole "ama", "osso", "esose" sono esempi di parole palindrome. Allora un programma capace di decidere se una parola x è palindroma o meno procederà al modo seguente:
1. codifica la parola x battuta alla tastiera in una sequenza di 0 d 1, cioè in un intero $cod(x)$
2. attiva un programma che a partire da tale numero dopo una serie di operazioni restituirà il valore 0 se x non è palindroma, il valore 1 se x è palindroma.

5. Problemi non numerici in matematica che nessun computer potrà mai risolvere.

Problemi di decidibilità e di indecibilità nascono anche in relazione ad insiemi non numerici. Ad esempio, sia L un linguaggio adeguato a parlare dei numeri interi. Nel prossimo capitolo si darà una descrizione precisa di tale linguaggio, per ora basti pensare che tale linguaggio deve contenere nel suo alfabeto i quantificatori $\forall$, $\exists$, i connettivi logici $\wedge$, $\vee$,$\neg$, le variabili x, y ... i nomi di operazioni + e $\cdot$ e le costanti 0 ed 1. E possibile allora chiedersi se l'insieme delle parole di tale linguaggio che esprimono asserzioni vere dell'aritmetica sia decidibile o meno.

Teorema 1. L'insieme delle asserzioni vere dell'aritmetica è indecidibile.

Dim. Per dimostrare tale teorema basta osservare che se fosse decidibile se una asserzione è vera o meno allora, in particolare, sarebbe decidibile se una asserzione del tipo

$$\exists x(a_h x^h + a_{h-1} x^{h-1} + ... + a_0 = 0)$$

sia vera o meno. Questo contrasta con il teorema di Matijasevic.

Integrazione e derivazione simbolica. Un altro risultato fondamentale di indecidibilità è legato al problema della integrazione elementare. È esperienza comune a tutti gli studenti di matematica che la derivazione e l'integrazione, a meno che non si lavori nell'ambito di un corso di calcolo numerico, sono procedimenti puramente formali. Si tratta in definitiva di trasformare delle "parole" rappresentanti funzioni in altre

"parole" rappresentanti le relative derivate o primitive. In questo caso si parlerà di derivazione e integrazione elementare o simbolica e si distingueranno tali procedimenti dall'integrazione numerica.
È possibile precisare tale modo di procedere formale al modo seguente. Consideriamo un alfabeto A che abbia tra i suoi simboli i nomi delle funzioni elementari, $A = \{+, \cdot, -, \backslash, x, sen, cos, log, \ldots\}$ più simboli per variabili e parentesi. Consideriamo poi il linguaggio $\mathcal{L} \subseteq A^+$ costituito dalle espressioni rappresentanti le funzioni che si ottengono per composizione dalle funzioni elementari. Per intenderci, appartengono ad $\mathcal{L}$ espressioni del tipo $sen(x\text{-}2)$, $(2x+1)/x$, $log(cos(x/(2x\text{-}x))$ mentre non appartengono ad $\mathcal{L}$ espressioni del tipo $sen(x,cos)x)$. Le regole per la derivazione che si studiano in un corso di analisi permettono di provare immediatamente la seguente proposizione.

Proposizione 2. L'operatore di derivazione $D : \mathcal{L} \to \mathcal{L}$ è una funzione totale computabile.

Dim. Basta osservare che D può essere calcolato, tramite un processo di ricorsione, tramite equazioni del tipo

$D(sen(x)) = cos(x)$
$D(sen(x)) = -sen(x)$
$D(log(x)) = 1/x$
$D(x^\lambda) = \lambda \cdot x^{\lambda\text{-}1}$
. . .

che rappresentano la "base" della ricorsione, ed equazioni del tipo

$D(f \cdot g) = f \cdot D(g) + g \cdot D(f)$
$D(f+g) = D(f)+D(g)$
. . .

che rappresentano il vero processo di ricorsione. □

Nel paragrafo 8 mostreremo un algoritmo scritto nel linguaggio *Mathematica* che permette di calcolare la derivata. Molto più complesso è il problema della integrazione. Naturalmente l'integrazione numerica non presenta problemi, essendo le funzioni rappresentate in $\mathcal{L}$ tutte continue ed esistendo numerosi metodi di calcolo numerico per il calcolo di un integrale. L'integrazione "simbolica", cioè la definizione di un procedimento che associa ad ogni elemento di $\mathcal{L}$ la relativa primitiva è cosa invece più complessa. Diremo che *una funzione in $\mathcal{L}$ è integrabile elementarmente* se ammette una primitiva in $\mathcal{L}$, cioè se appartiene al codominio di $D : \mathcal{L} \to \mathcal{L}$. Esistono funzioni in $\mathcal{L}$ che non sono integrabili elementarmente. Vale inoltre il seguente teorema che prova la non esistenza di un procedimento generale capace di dire se una funzione è integrabile elementarmente o meno.

Teorema 3. L'insieme delle funzioni elementari integrabili elementarmente è effettivamente enumerabile ma non decidibile. Pertanto non esiste un algoritmo che riceve in input una funzione elementare e che:

- restituisca una sua primitiva se la funzione è integrabile elementarmente
- avverta se la funzione non è integrabile elementarmente.

Dim (parte facile). Ci limitiamo ad osservare che l'insieme delle funzioni integrabili elementarmente è effettivamente enumerabile in quanto codominio dell'operatore di derivazione D. □

Il fatto che l'insieme delle funzioni integrabili elementarmente è parzialmente decidibile significa che un processo di calcolo della primitiva di una funzione data f potrebbe consistere nel

- produrre in successione tutti gli elementi di $\mathcal{L}$,
  $f_1, f_2, f_3, \ldots, f_i, \ldots$
- man mano calcolare le derivate di tali funzioni:
  $f_1', f_2', f_3', \ldots, f_i' \ldots$
- verificare se f_i' coincide con f
- in caso positivo dare come risultato f_i e fermarsi, altrimenti proseguire.

Naturalmente se f non è integrabile elementarmente un tale algoritmo non si ferma mai. Un tale modo di procedere consiste in un certo senso nel "procedere per tentativi" in maniera sistematica in modo da esplorare tutte le possibilità.

6. Operatori algebrici.
Prima di parlare dei sistemi di riscrittura, è opportuno ricordare alcune nozioni riguardanti gli operatori algebrici e le relazioni.

Definizione 1. Sia S un insieme e denotiamo con $\mathcal{P}(S)$ l'insieme delle sua parti, allora chiameremo *operatore algebrico su S* ogni funzione $T : \mathcal{P}(S) \to \mathcal{P}(S)$ tale che:

a) $T(X) \supseteq X$;
b) $X \subseteq Y \Rightarrow T(X) \subseteq T(Y)$;
c) $x \in T(X) \Rightarrow \exists F \subseteq X$, F finito, $x \in T(F)$.

Diremo che T è di *tipo k* se la cardinalità dell'insieme F di cui è detto in c) è sempre minore di k. La teoria degli operatori algebrici deve essere vista come un modo astratto di considerare un processo con cui si costruiscono nuovi oggetti a partire da un dato insieme X di oggetti. Allora
a) significa che tra le cose che posso costruire con X ci sono gli elementi di X,
b) significa che se una cosa può essere costruita a partire da X allora (a maggior ragione) può essere costruita a partire da un insieme che contiene X
c) afferma che dire che una cosa è costruibile a partire da X significa in realtà che è costruibile a partire da un numero finito di elementi di X. In altre parole, significa che il processo di costruzione è finitario.

Definizione 2. Diremo che un operatore $D : \mathcal{P}(S)\to\mathcal{P}(S)$ è un *operatore di chiusura* se verifica le condizioni a), b) e

d) $D(D(X)) \subseteq D(X)$,

Un *operatore di chiusura algebrico* è un operatore che verifica a), b), c) e d).

La condizione d) si può interpretare dicendo che se un oggetto x è costruito con materiale in $D(X)$ il quale a sua volta è stato costruito con materiale in X allora tale oggetto è, di fatto, costruito con materiale in X. In altre parole a partire dagli oggetti in $D(X)$ non è possibile costruire niente che non sia già in $D(X)$. Si noti che, poiché per a) si ha che $D(D(X))=D(X)$, se D è un operatore di chiusura algebrico allora risulta che $D(D(X))=D(X)$.

Definizione 3. Dato un operatore T, si chiama *punto fisso* o *punto unito* di T un insieme X tale che $T(X) = X$.

Se T è un operatore algebrico allora X è un punto fisso se e solo se $T(X)\subseteq X$ essendo l'inclusione $T(X)\supseteq X$ sempre verificata. Allora i punti fissi sono gli insiemi "saturi" rispetto al processo di costruzione T, cioè insiemi da cui non è possibile ottenere nuovo materiale utilizzando T. Se T è un operatore di chiusura allora, essendo $T(T(X)) = T(X)$, ogni $T(X)$ risulta essere un punto fisso.
Nel prossimo teorema scriveremo $T^n(X)$ per indicare il risultato della applicazione dell'operatore T n volte; più precisamente definiamo $T^n(X)$ per ricorsione su n tramite le equazioni
- $T^0(X) = X$
– $T^{n+1}(X) = T(T^n(X))$.

Teorema 4. Sia $T : \mathcal{P}(S)\to\mathcal{P}(S)$ un operatore algebrico ed X un insieme, allora l'insieme
$$D(X) = \cup_{n\in N}T^n(X) \tag{3}$$
è il minimo punto fisso di T contenente X.

Dim. Per provare che $D(X)=\cup T^n(X)$, cominciamo con il provare che $\cup T^n(X)$ è un punto unito e quindi che $T(\cup_{n\in N} T^n(X))\subseteq \cup_{n\in N} T^n(X)$. Ora da $T(X)\supseteq X$, applicando n volte l'operatore T, segue che $T^{n+1}(X)\supseteq T^n(X)$. Ciò prova che $T^n(X)$ è una successione crescente. Sia x un elemento di $T(\cup_{n\in N} T^n(X))$, allora essendo T algebrico risulterà che $x\in T(F)$ per F opportuno sottoinsieme finito di $\cup_{n\in N} T^n(X)$. Per la finitezza di F e la crescenza di $T^n(X)$ esisterà $p\in N$ tale che F è contenuto in $T^p(X)$. Ne segue che

$$x\in T(F) \subseteq T(T^p(X))=T^{p+1}(X) \subseteq \cup_{n\in N} T^n(X).$$

Ciò prova che $T(\cup_{n\in N} T^n(X))$ è contenuto in $\cup_{n\in N} T^n(X)$ e che quindi $\cup_{n\in N} T^n(X)$ è un punto unito. Per provare che $\cup_{n\in N} T^n(X)$ è il minimo punto unito contenente X, sia M un qualsiasi punto unito contenente X, allora è immediato che $T^n(M) = M$. D'altra parte, applicando n volte l'operatore T alla diseguaglianza $M\supseteq X$, otteniamo che $T^n(M) \supseteq T^n(X)$ e quindi che $M \supseteq T^n(X)$. In definitiva M contiene $\cup_{n\in N} T^n(X)$ e quindi $\cup_{n\in N} T^n(X)$ è contenuto in ogni punto unito M contenente X. □

Esercizio. Sia R l'insieme dei numeri reali e $T : \mathcal{P}(R)\rightarrow \mathcal{P}(R)$ l'operatore definito ponendo

$$T(X)=X\cup\{3x+x' \mid x\in X, x'\in X\}.$$

Provare che T è un operatore algebrico e trovare il minimo punto fisso contenente $\{2,5\}$.

Esempio in algebra. Sia G un gruppo moltiplicativo. Allora,dati due sottoinsiemi X ed Y di G, poniamo $X\cdot Y =\{x\cdot y : x\in X, y\in Y\}$. In altre parole, indichiamo con $X\cdot Y$ l'insieme degli elementi che si possono ottenere come prodotto di un elemento di X per un elemento di Y. In particolare, poniamo $X\cdot X = \{x_1\cdot x_2 : x_1\in X, x_2\in X\}$. Inoltre poniamo $X^{-1} = \{x^{-1} : x\in X\}$. Allora è possibile definire un operatore $T : \mathcal{P}(R)\rightarrow \mathcal{P}(R)$ ponendo:

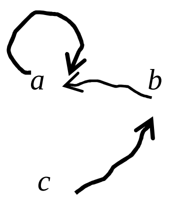
$$T(X) = X\cup X\cdot X\cup X^{-1}\cup\{1\}.$$

Un insieme X è un punto fisso di T se e solo se $T(X)\subseteq X$ cioe se e solo se $X\cdot X\subseteq X$, $X^{-1}\subseteq X$, $1\in X$. Questo significa che X è un punto unito se e solo se è chiuso rispetto al prodotto, all'inverso e se contiene 1. In altri termini i punti fissi di T coincidono con i sottogruppi di G. Non è difficile vedere che T è un operatore algebrico. Pertanto, dato un sottoinsieme X di G, $\cup_{n\in N} T^n(X)$ è il minimo punto fisso di T contenente X, cioè è il più piccolo sottogruppo di G contenente X. Tale sottogruppo viene di solito chiamato *sottogruppo generato da X*

7. Calcolo delle relazioni

Definizione 1. Siano $D_1,...,D_n$ insiemi, allora chiamiamo *relazione n-aria* un sottoinsieme $\mathcal{R}$ del prodotto cartesiano $D_1\times ...\times D_n$. In particolare, se D è un insieme, chiamiamo *relazione n-aria su D* un sottoinsieme del prodotto cartesiano D^n.

Per le relazioni binarie si usa in generale la *notazione infissa*, cioè si scrive $x\mathcal{R}y$ per denotare che $(x,y)\in \mathcal{R}$, cioè per denotare che x ed y sono nella relazione $\mathcal{R}$. Ad esempio, nel caso di relazione d'ordine $\leq$, scriveremo $x\leq y$ al posto di $(x,y)\in\leq$. Un insieme D con una relazione binaria viene anche detta *grafo*. Un grafo si rappresenta sul piano euclideo fissando per ogni elemento di D un punto e tracciando, per ogni coppia di punti p_1 e p_2 una freccia che parte da p_1 e raggiunge p_2 ogni volta che la coppia (p_1,p_2) corrisponde a due elementi di D che stanno nella relazione $\mathcal{R}$. Ad esempio, se $D = \{a,b,c\}$ ed $\mathcal{R} = \{(a,a),(b,a),(c,b)\}$, allora disegniamo tre punti in corrispondenza di a, b e c e rappresentiamo il grafo al modo seguente:



Elenchiamo ora le principali proprietà che una relazione può soddisfare.

Definizione 2. Sia $\mathcal{R}$ una relazione binaria su di un insieme D, allora diciamo che

- $\mathcal{R}$ è *riflessiva* se $x\mathcal{R}x$ per ogni $x\in D$
- $\mathcal{R}$ è *simmetrica* se $x\mathcal{R}y \Rightarrow y\mathcal{R}x$ per ogni x ed y in D
- $\mathcal{R}$ è *transitiva* se $x\mathcal{R}y$ e $y\mathcal{R}z \Rightarrow x\mathcal{R}z$ per ogni $x,y,z\in D$
- $\mathcal{R}$ è *asimmetrica* se $x\mathcal{R}y$ e $y\mathcal{R}x \Rightarrow x=y$ per ogni x ed y in D
- $\mathcal{R}$ è *antisimmetrica* se $x\mathcal{R}y \Rightarrow$ non $y\mathcal{R}x$ per ogni x ed y in D

- $\mathcal{R}$ è *totale* o *lineare* se $x\mathcal{R}y$ oppure $y\mathcal{R}x$ per ogni x ed y in D.

Le seguenti notazioni premettono di esprimere in maniera più elegante le nozioni ora date.

Definizione 3. La relazione *identità* è la relazione $Diag(D) \subseteq D^2$ è definita da

$$Diag(D)=\{(x,x) \mid x\in D\},$$

e di solito viene indicata con =. Date due relazioni $\mathcal{R}_1 \subseteq X\times Y$ e $\mathcal{R}_2 \subseteq Y\times Z$ chiamiamo *composizione* di $\mathcal{R}_1$ e $\mathcal{R}_2$ la relazione $\mathcal{R}_1\circ \mathcal{R}_2 \subseteq X\times Z$ definita ponendo

$$\mathcal{R}_1\circ \mathcal{R}_2 =\{(x,z) \mid \exists y\in Y \text{ tale che } (x\mathcal{R}_1 y \text{ e } y\mathcal{R}_2 z\}.$$

Data una relazione binaria $\mathcal{R}\subseteq X\times Y$ chiamiamo *inversa* di $\mathcal{R}$ la relazione binaria $\mathcal{R}^{-1}\subseteq Y\times X$ definita ponendo

$$\mathcal{R}^{-1}=\{(y,x) \mid x\mathcal{R}y\}.$$

La relazione $Diag(D)$ viene anche detta anche *diagonale di D*.

Proposizione 4. Sia $\mathcal{R}$ una relazione binaria su D, allora:

i) $\mathcal{R}$ riflessiva $\Leftrightarrow$ $\mathcal{R} \supseteq Diag(D)$
ii) $\mathcal{R}$ simmetrica $\Leftrightarrow$ $\mathcal{R} = \mathcal{R}^{-1}$
iii) $\mathcal{R}$ transitiva $\Leftrightarrow$ $\mathcal{R}\circ\mathcal{R} \subseteq \mathcal{R}$
iv) $\mathcal{R}$ asimmetrica $\Leftrightarrow$ $\mathcal{R}\cap\mathcal{R}^{-1}\subseteq Diag(D)$
v) $\mathcal{R}$ antisimmetrica $\Leftrightarrow$ $\mathcal{R}\cap\mathcal{R}^{-1} = \emptyset$
vi) $\mathcal{R}$ totale se $\mathcal{R}\cup\mathcal{R}^{-1} = D\times D$.

Dim. Si lascia come esercizio. □

Proposizione 5. Poniamo

$$Rifl(\mathcal{R}) = \mathcal{R}\cup Diag(D).$$

Allora $Rifl : \mathcal{P}(D^2)\to\mathcal{P}(D^2)$ è un operatore di chiusura i cui punti fissi sono le relazioni riflessive. Pertanto, $Rifl(\mathcal{R})$ e la più piccola relazione riflessiva contenente $\mathcal{R}$.

Ad esempio se < e l'usuale relazione di ordine stretto in N allora $Rifl(<)$ è la relazione ≤.

Proposizione 6. Poniamo

$$Simm(\mathcal{R}) = \mathcal{R}\cup\mathcal{R}^{-1}.$$

Allora $Simm : \mathcal{P}(D^2)\to\mathcal{P}(D^2)$ è un operatore di chiusura i cui punti fissi sono le relazioni simmetriche. Pertanto $Simm(\mathcal{R})$ è la più piccola relazione simmetrica contenente $\mathcal{R}$.

Proposizione 7. Poniamo

$$Tr(\mathcal{R}) = (\mathcal{R}\circ\mathcal{R})\cup\mathcal{R}$$

Allora $Tr : \mathcal{P}(D^2)\to\mathcal{P}(D^2)$ è un operatore algebrico i cui punti fissi sono le relazioni transitive. Ne segue che l'operatore $Trans : \mathcal{P}(D^2)\to\mathcal{P}(D^2)$ definito ponendo

$$Trans(\mathcal{R}) = \cup Tr^n(\mathcal{R}),$$

è un operatore di chiusura algebrico i cui punti fissi sono le relazioni transitive e $Trans(\mathcal{R})$ è la più piccola relazione transitiva contenente $\mathcal{R}$.

Problema. Sia $\mathcal{R}$ una relazione binaria e chiamiamo *percorso* una successione $x_1, ...,x_n$ tale che $x_i\mathcal{R}x_{i+1}$ per i = 1,...,n-1. Poniamo inoltre

$$\mathcal{R}' = \{(x,y) \mid \text{esiste un percorso } x_1,...,x_n \text{ tale che } x_1 = x,\ x_n = y\}.$$

Dimostrare che $\mathcal{R}' = Trans(\mathcal{R})$.

Tra le relazioni binarie rivestono particolare importanza le relazioni d'ordine e le relazioni di equivalenza. Ricordiamo che una relazione binaria è chiamata:
- preordine se è riflessiva e transitiva
- ordine se è una relazione di preordine asimmetrica
- ordine stretto se è una relazione transitiva e antisimmetrica
- ordine lineare se è una relazione di ordine totale.

Proposizione 8. Poniamo

$$Ord(\mathcal{R}) = Trans(Rifl(\mathcal{R})).$$

Allora $Ord : \mathcal{P}(D^2)\to\mathcal{P}(D^2)$ è un operatore di chiusura algebrico i cui punti fissi sono le relazioni di preordine e $Ord(\mathcal{R})$ è la più piccola relazione di preordine contenente $\mathcal{R}$.

Una relazione binaria $\mathcal{R}$ su di un insieme D è chiamata *equivalenza* se è riflessiva, simmetrica e transitiva, cioè se:

$$\mathcal{R} \supseteq Diag(D) \ ; \ \mathcal{R} = \mathcal{R}^{-1} \ ; \ \mathcal{R}\circ\mathcal{R}\subseteq\mathcal{R}.$$

Una relazione di equivalenza si indica in genere con il simbolo $\equiv$. L'importanza della nozione di relazione di equivalenza deriva dal fatto che è strettamente legata al processo di astrazione cioè al processo per cui si astrae da una particolare proprietà degli oggetti che si esaminano. Ciò comporta che vengono considerati "praticamente uguali", cioè equivalenti, due oggetti che si distinguono solo per tale proprietà. Ad esempio se non si considerano importanti le dimensioni di una fotografia ma solo le immagini che essa contiene, allora il considerare equivalenti una foto ed un suo ingrandimento equivale a fare astrazione dalle dimensioni. Se ci riferiamo all'insieme dei triangoli e non si interessa la particolare posizione occupata da un triangolo, allora saremo portati a considerare equivalenti due triangoli che siano sovrapponibili tramite un movimento.

Problema. Nell'insieme dei numeri reali poniamo $x\mathcal{R}_1y$ se x differisce da y per meno di un decimo. Poniamo poi $x\mathcal{R}_2y$ se x ed y hanno la stessa parte intera. Dire se le due relazioni ora definite sono relazioni di equivalenza.

Proposizione 9. Poniamo

$$Eq(\mathcal{R}) = Trans(Simm(Rifl(\mathcal{R}))).$$

Allora $Eq : \mathcal{P}(D^2)\to\mathcal{P}(D^2)$ è un operatore di chiusura algebrico i cui punti fissi sono le relazioni di equivalenza. Ne segue che $Eq(\mathcal{R})$ è la più piccola relazione di equivalenza contenente $\mathcal{R}$.

La nozione di equivalenza è strettamente collegata a quella di partizione. Una classe Π di sottoinsiemi di S è detta una *partizione* di S se:
- gli elementi di Π sono a due a due disgiunti,
- l'unione degli elementi di Π coincide con S.

Gli elementi di Π vengono anche chiamati *classi*. In altre parole Π è una partizione se rappresenta un modo di dividere gli elementi di S in tante parti separate. Ad esempio se S è l'insieme dei numeri interi, P è l'insieme dei numeri pari e D è l'insieme dei numeri dispari, allora una partizione di S si ottiene ponendo $\Pi = \{P,D\}$.

Proposizione 10. Sia $\equiv$ una relazione di equivalenza in un insieme S ed indichiamo, per ogni $z \in S$, con [z] la classe $\{z'\in S \mid z' \equiv z\}$, cioè l'insieme degli elementi equivalenti a z. Allora l'insieme $\Pi = \{[z] \mid z\in S\}$ è una partizione. Viceversa, sia Π una partizione di S e definiamo la relazione binaria $\equiv$ ponendo $x \equiv y$ se e solo se x ed y appartengono allo stesso elemento in Π. Allora $\equiv$ è una relazione di equivalenza.

Dim. Supponiamo che $\equiv$ sia una relazione di equivalenza. Per la proprietà riflessiva $z\in[z]$ e ciò prova che ogni elemento di S appartiene ad una classe in Π. Se poi $[z]$ e $[z']$ sono due classi e $[z]\neq[z']$, allora non può esistere un elemento x comune a tali classi. Infatti in tale caso risulterebbe che $z\equiv x$ e $z'\equiv x$ e quindi per la proprietà simmetrica $z\equiv x$ e $x\equiv z'$ e quindi per la proprietà transitiva $z \equiv z'$ (in contrasto con l'ipotesi $[z] \neq [z']$).

Sia Π una partizione. Che $\equiv$ sia riflessiva deriva dal fatto che per ogni x esiste $X\in \Pi$ tale che $x\in X$. Pertanto $x\equiv x$. La proprietà simmetrica è immediata. Per provare la proprietà transitiva supponiamo che $x\equiv y$ e $y\equiv z$, cioè che esistono X_1 ed X_2 in Π tali che $x,y\in X_1$ e $y, z\in X_2$. Allora X_1 ed X_2 hanno in comune l'elemento y. Poiché due elementi distinti di Π sono disgiunti, ciò comporta che $X_1=X_2$ è quindi che x e z appartengono ad uno stesso elemento di Π. Allora $x\equiv z$ e pertanto $\equiv$ è una relazione di equivalenza. □

Ogni classe $[z]$ è chiamata *classe completa di equivalenza* e z viene chiamato *elemento rappresentativo* di tale classe, la partizione Π è chiamata *quoziente di S modulo* $\equiv$ e si indica con $S/\equiv$.

8. Giochi, labirinti, solitari e sistemi di riscrittura.

In questo paragrafo introdurremo strutture matematiche che permettono di rappresentare le più svariate situazioni in cui si debba risolvere un problema. Cominciamo con l'osservare che spesso

<u>*risolvere un problema consiste nel fare una serie di "azioni" che facciano passare da uno stato iniziale ad uno stato che si possa considerare soluzione del problema.*</u>

Naturalmente non tutte le azioni sono possibili e questo perché alcune o sono vietate oppure non sono "fisicamente" possibili. Allora nell'insieme S degli stati è definita una relazione binaria R il cui significato è che $(x,y)\in R$ se e solo se è possibile passare dallo stato x allo stato y. Consideriamo ad esempio un qualunque solitario di carte in cui si dispongono in qualche modo le carte sul tavolo, Chiamiamo *stato* una qualunque possibile distribuzione di carte sul tavolo ed indichiamo con S l'insieme degli stati. Si risolve il solitario se si raggiunge una configurazione c che viene considerata vincente (ad esempio disporre le carte in ordine crescente in quattro gruppi dello stesso colore). Le regole del solitario si possono rappresentare da una relazione binaria $R \subseteq S\times S$ in modo che sRs' che esiste una mossa (corretta in base al regolamento del gioco) che permette di passare dallo stato s allo stato s'. Risolvere il solitario significa trovare una successione $s_1,s_2,...s_n$ di stati tali che

i) s_1 sia lo stato iniziale (in generale scelto in modo casuale cioè dopo aver mescolato le carte)
ii) $(s_i,s_{i+1})\in R$ cioè è lecito passare da s_i a s_{i+1} per ogni $i = 1,...,n-1$
iii) s_n appartenga all'insieme delle configurazioni considerate vincenti.

Spesso la relazione R viene rappresentata tramite una freccia $\rightarrow$ e si scrive $x\rightarrow y$ per indicare che è possibile passare da x ad y. In definitiva abbiamo la seguente definizione:

Definizione 8.1. Chiamiamo *sistema di riduzione* o *sistema di trasformazioni* una coppia $(S,\rightarrow)$ con S insieme non vuoto e $\rightarrow$ relazione binaria su S. Gli elementi di S vengono chiamati *stati*. Chiameremo *solitario* o *gioco ad una persona* una struttura $(S,\rightarrow,GOAL)$ dove $(S,\rightarrow)$ è un sistema di trasformazioni e $GOAL$ è un insieme di stati chiamati *stati vincenti*.

Se la coppia x e y di elementi di S è nella relazione $\rightarrow$ allora diremo che *y è un ridotto diretto di x* e scriveremo $x\rightarrow y$. Chiamiamo *successione di riduzioni* una successione $a_1,...,a_n$ di elementi di S tale che ogni a_i è il ridotto diretto di a_{i-1}.

Definizione 8.2. Dato un gioco $(S,\rightarrow,GOAL)$ chiamiamo *soluzione di stato iniziale a* una successione di riduzioni $a_1,...,a_n$ tale che $a_1=a$ ed $a_n\in GOAL$. Chiamiamo *strategia* una funzione $f : S\rightarrow S$ che associa ad ogni stato $x\in S$ un nuovo stato $f(x)$ tale che $x\rightarrow f(x)$ se $\{y \in S : x\rightarrow y\}$ è non vuoto, $f(x) = x$ altrimenti. Una strategia si dice *vincente* se per ogni x esiste $n\in N$ tale che la successione $x, f(x), f^2(x), ...,f^n(x)$ è una soluzione di stato iniziale x.

Pertanto una strategia vincente permette, dato un qualunque stato di partenza x, di arrivare ad uno stato vincente $f^n(x)$ dopo un numero finito n di "mosse".

Un labirinto

Per mostrare come un tale formalismo permetta di rappresentare le situazioni più diverse, consideriamo uno dei problemi più antichi, quello dei labirinti. Un labirinto è un percorso in cui di tanto in tanto si deve scegliere tra più strade quale strada seguire. Lo scopo del gioco è quello di uscire dal labirinto. Un modo equivalente di rappresentare un labirinto è quello di considerare un insieme S di “stanze” più una relazione $\rightarrow$ definita ponendo $x \rightarrow y$ se esiste un collegamento (una porta) tra una stanza e l’altra. Alcune stanze sono considerate “vincenti” (ad esempio le stanze che comunicando con l’esterno e che permettono quindi di uscire, oppure stanze in cui esiste un “tesoro”). Risolvere il gioco del labirinto a partire da una data stanza s_1 significa trovare una successione $s_1, s_2, \ldots s_n$ di stanze tale che:

i) per ogni $i = 1,\ldots,n\text{-}1$ sia possibile passare da s_i a s_{i+1}
iii) s_n sia una delle stanze in cui esiste una porta di uscita.

Un altro famoso gioco è il gioco del quindici. Si tratta di un quadrato di plastica su cui sono inserite 15 tessere quadrate numerate mentre un quadrato rimane vuoto. Le mosse consentite sono solo spostare una delle tessere nello spazio vuoto (liberando in questo modo un altro spazio vuoto). Si vince se si riesce, partendo da una configurazione iniziale scelta a caso, a raggiungere una configurazione che si ritiene vincente. Ad esempio può essere considerata vincente la configurazione in cui tutti i numeri sono messi successivamente come nella figura affianco.

Il gioco del quindici

Una categoria di giochi che sono molto collegati con la matematica sono i *sistemi di riscrittura* che sono giochi i cui gli stati sono parole di un linguaggio. Ad esempio consideriamo il problema di risolvere le equazioni di primo grado in una incognita. Chiamiamo con S l'insieme di tutte le possibile equazioni di primo grado, ad esempio

$$3x+5 = x\text{-}3x, \ \ x = x\text{-}3(x+1), \ \ . . .$$

Poniamo *GOAL* uguale all’insieme delle equazioni del tipo $x = k$. Se chiamiamo equivalenti due equazioni che abbiano le stesse soluzioni, il problema che si pone è, data una equazione, di arrivare ad una equazione equivalente appartenente a *Goal*. Ciò deve essere fatto "rispettando le regole", cioè tramite una serie di operazioni che facciano passare da una equazione ad una equazione equivalente, cioè che abbia le stesse soluzioni.
Le operazioni che usualmente si utilizzano sono le seguenti:
1. passare una quantità che si somma da un lato all'altro di una equazione facendola diventare una quantità che si sottrae
2. passare una quantità che si sottrae da un lato all'altro di una equazione facendola diventare una quantità che si somma
3. passare una quantità che si moltiplica da un lato all'altro facendola diventare una quantità che divide
4. passare una quantità che si divide da un lato all'altro facendola diventare una quantità che moltiplica
5. applicare tute le leggi dell'aritmetica (proprietà distributiva, proprietà commutativa, ...)
6. effettuare tutti i calcoli che si possono fare.

Ad esempio, partendo da $3x+5=x\text{-}3x+1$ e saltando qualche passaggio, otteniamo

$3x+5=x\text{-}3x+1$ (punto di partenza)
$3x+5 = \text{-}2x+1$ (per la regola 6)
$3x+5+2x = +1$ (per la regola 2)
$5x +5= +1$ (per la regola 6)
$5x = \text{-}5+1$ (per la regola 1)
$5x = \text{-}4$ (per la regola 6)
$x = \text{-}4/5$ (per la regola 3)

In questo esempio una strategia consiste nello scegliere ad ogni passo nell’ordine:
“di fare i calcoli che si possono fare”
“di spostare tutte le variabili al primo membro dell’equazione”

"di spostare tutte le costanti al secondo membro dell'equazione"

Definizione 8.3. Un *sistema di riscrittura* è un gioco (S,$\rightarrow$,$GOAL$) in cui S è l'insieme delle parole in un linguaggio L. Un sistema di *riduzione a forma normale* è un sistema di riscrittura in cui $GOAL = \{n \in S$: non esiste $b \neq n$ tale che $n \rightarrow b\}$. In tale caso diciamo che un elemento $n \in GOAL$ è in *forma normale.*

Da notare che nell'esempio delle equazioni abbiamo specificato anche la regola adottata in ogni passo. Questo significa che il gioco è stato descritto fornendo più regole ed una soluzione del gioco consiste anche nello specificare quale regola si è applicata. Quando si procede in questo modo è più utile la seguente definizione.

Definizione 8.4. Chiamiamo *sistema di riscrittura a più regole* una struttura (S, $(\rightarrow_i)_{i\in I}$, $GOAL$) dove S è un linguaggio su di un dato alfabeto e $(\rightarrow_i)_{i\in I}$ è una famiglia di relazioni binarie dette *regole di riscrittura*.

Come viene fatto di solito per le relazioni binarie useremo la notazione infissa scrivendo $x\rightarrow_i y$ per indicare che $(x,y)\in \rightarrow_i$ cioè che x è nella relazione $\rightarrow_i$ con y. In tale caso diremo anche che la parola x si può *riscrivere nella parola y in accordo con la regola di indice i.*

Proposizione 8.5. Dato un sistema di riscrittura a più regole (S, $(\rightarrow_i)_{i\in I}$, $GOAL$), viene definito un sistema di riscrittura ad una regola (S, $\rightarrow$, $GOAL$), ponendo $\rightarrow$ uguale alla relazione che si ottiene come unione delle relazioni $\rightarrow_i$.

In altri termini poniamo $x\rightarrow y$ se esiste $i\in I$ tale che $x\rightarrow_i y$. Ci riferiamo a $\rightarrow$ quando non interessa sapere quale regola è stata utilizzata per passare dallo stato x allo stato y ma solo che è consentito passare da x ad y.

Esempio, espressioni aritmetiche. Sia L l'insieme delle espressioni che si possono scrivere coinvolgendo i numeri interi, la somma, il prodotto e la divisione. Ad esempio avremo in L espressioni del tipo

$$(3+5)\cdot 37+1,(1/3+1)/2, (2/3)\cdot(3/4+1), \ldots.$$

Possiamo allora chiamare *forma normale* ogni espressione del tipo n/m con n ed m rappresentazioni decimali di due interi primi tra loro. Allora quando si esegue una operazione, ad esempio il prodotto, il problema consiste nel passare dalla forma normale di due numeri x ed y alla forma normale di $x\cdot y$. Ad esempio, la somma di 3/5 più 2/3 seguirà la procedura di riscrittura

(3/5)+(1/3) (punto di partenza)
(3·3+5·1)/(5·3) (per la regola della somma per cui $n/m+p/q$ si può riscrivere in $(n\cdot q+m\cdot p)/(m\cdot q)$)
(19+10)/15 (per le regole di calcolo del prodotto di due interi)
29/15 (per le regole di calcolo della somma di due interi)

Il prodotto di 3/5 per 2/3 avviene secondo le seguenti riscrittura
(3/5)·(1/3) (punto di partenza)
(3·1)/(5·3) (per la regola per cui $n/m\cdot p/q$ si può riscrivere in $(n\cdot p)/(m\cdot q)$)
3/15 (per le regole di calcolo del prodotto)
1/5 (per le regole che permettono la riduzione di una frazione a numeratore e denominatore primi tra loro)

In generale si presenta la seguente situazione. Dato un linguaggio L si considera una relazione di equivalenza in L che, detto in termini intuitivi, corrisponde all'idea che due espressioni in L "rappresentano la stessa cosa". Ad esempio:
- nel calcolo proposizionale sono *logicamente equivalenti* due formule che danno luogo alla stessa tavola di verità
- in aritmetica consideriamo equivalenti due espressioni che denotano lo stesso numero
- in algebra chiamiamo equivalenti due equazioni che hanno le stesse soluzioni.

Appare allora naturale porsi il problema di come si possa trasformare una espressione di L in una più semplice che sia equivalente. Questo viene fatto appunto tramite opportune "regole di riscrittura" in cui possiamo interpretare $x\rightarrow y$ dicendo che y è equivalente ad y ma è "più semplice" di x. Pertanto una parola in forma normale è una parola che non può essere ulteriormente semplificata. Il gioco consiste allora nel partire

da una parola x e trasformarla fino a giungere ad una parola che non può essere più semplificata cioè che sia in forma normale.
Per indicare una tale successione scriveremo anche $a_1 \to a_2 ... \to a_{n-1} \to a_n$. Analogamente, x è convertibile in y se e solo $x = y$ oppure se esiste una successione $a_1,...,a_n$ tale che $a_1 \leftrightarrow a_2$, $a_2 \leftrightarrow a_3,..., a_{n-1} \leftrightarrow a_n$.

Definizione 8.6. Dato un gioco (S, $\to$, $GOAL$), indichiamo con $\longrightarrow$ la chiusura transitiva di $\to$. Indichiamo invece con $\leftrightarrow$ la chiusura simmetrica di $\to$ e con $\longleftrightarrow$ la relazione di equivalenza generata da $\to$.

Siano x ed y elementi di S, allora :
- se $x \longrightarrow y$ diremo che y è *un ridotto* di x
- se $x \longleftrightarrow y$ diremo anche che x ed y sono *convertibili*.

Problema. Un problema di notevole importanza è il seguente:
Dato un insieme S ed una relazione di equivalenza $\equiv$ trovare un sistema di riduzione la cui relazione di equivalenza associata $\longleftrightarrow$ coincida con $\equiv$ ed inoltre tale che ogni elemento può essere ridotto a forma normale unica.

L'importanza di questo problema è che se si riesce a risolverlo allora in ogni classe di equivalenza è possibile trovare un particolare elemento rappresentativo ed inoltre è possibile lavorare sugli elementi rappresentativi invece che sulle classi.

9. Riscrittura come programmazione: il calcolo simbolico (Mathematica e Derive).
La nozione di sistema di riscrittura è alla base di linguaggi di programmazione che, come *Mathematica*, permettono di effettuare *"calcolo simbolico"*. Esporremo alcune idee generali di Mathematica in maniera piuttosto approssimativa (il manuale di Mathematica è un libro enorme che pochi riescono a leggere tutto !).
Indichiamo con A l'insieme dei caratteri della tastiera che non abbiano un significato "speciale". Questo significa in A non mettiamo alcuni caratteri come * (che denota il prodotto), + (che denota la somma), ^ (che denota la funzione potenza) ed altri caratteri simili.

Definizione 1. L'insieme dei *termini* in Mathematica è il linguaggio generato dalla grammatica il cui alfabeto terminale è l'insieme A dei caratteri (non speciali) della tastiera, il cui alfabeto ausiliario è l'insieme {*nome_di_funzione*, *costante*, *variabile*, *termine*}, il cui *start-symbol* è *"termine"* ed avente come insieme di produzioni:
- *nome_di_ funzione* $\to p$ (con $p \in A^+$)
- *costante* $\to p$ (con $p \in A^+$)
- *variabile* $\to$ *p_* (con $p \in A^+$)
- *termine* $\to$ *costante*
- *termine* $\to$ *variabile*
- *termine* $\to$ *nome_di_funzione*[*termine*,...,*termine*].

In termini più semplici il linguaggio dei termini in Mathematica è definito dalle seguenti regole:
- un *nome di funzione* è una qualunque parola in A
- una *costante* è una qualunque parola in A
- una *variabile* è una qualunque parola che termini con il simbolo _ (detto *underscore*)
- ogni costante o variabile è un *termine*
- se $t_1,...,t_n$ sono termini ed f è un nome di funzione allora $f[t_1,...,t_n]$ è un termine[2]

L'idea intuitiva è che un termine sia un modo per descrivere una funzione. Da notare che Mathematica preferisce, differentemente dalla pratica comune, l'uso delle parentesi quadrate invece che le tonde nella

[2] Una tale notazione è pertanto di tipo *prefisso*, cioè il nome della funzione precede i valori a cui è applicata. Tuttavia Mathematica accetta anche modi di scrivere che sono più vicine alle abitudine dei matematici. Ad esempio permette di scrivere 2+3 e quindi di utilizzare la notazione infissa. Tuttavia il modo in cui Mathematica immagazzina in memoria ed elabora tale termine è del tipo Plus[2,3].

costruzione dei termini. Tenendo conto che un termine si dice *chiuso* se non contiene variabili, i seguenti sono esempi sono termini chiusi

abc, ss_, cc[a], wws[ax,cc[a]],

Sono invece termini nelle variabili x_ ed y_

Times[x_,y_], Plus[coseno[x_],5], . . .

La nozione di *sottotermine* si ottiene assumendo che tale relazione sia transitiva, riflessiva e che $t_1,...,t_n$ siano sottotermini di $f[t_1,...,t_n]$. Allora, per esempio, sono sottotermini del termine

Times[*f*[3],*Cos*[Times[2,x_]]]

i seguenti termini:

Times[*f*[3],*Cos*[*Times*[2,x_]]], *f*[3], *Cos*[*Times*[2,x_]], 3, *Times*[2,*x*_], 2, x_.

Se indichiamo con $t(x_1,...,x_n)$ un termine in cui compaiono le variabili libere x_1_,...,x_n_ e se $t_1,...,t_n$sono termini, allora indichiamo con $t(x_1/t_1,...,x_n/t_n)$ il termine che si ottiene sostituendo in $t(x_1,...,x_n)$ al posto delle variabili x_1_,...,x_n_ i termini $t_1,...,t_n$.

Definizione 2. Chiamiamo *regola di riscrittura in Mathematica* una espressione del tipo

$$ter_1(x_1,...,x_n) := ter_2(x_1,...,x_n)$$

dove ter_1 e te_2 sono due termini che prendono il nome di *testa* e *corpo* della regola. Un *programma in Mathematica* è un insieme di regole di riscrittura in cui due regole non hanno mai la stessa testa.

Se indichiamo con *TC* l'insieme dei termini chiusi un programma determina un sistema di riscrittura in *TC* al modo seguente:

- siano $t_1,...,t_n$ termini chiusi, allora ogni termine contenente come sottotermine $ter_1(x_1/t_1,...,x_n/t_n)$ può essere riscritto sostituendo al posto di $ter_1(x_1/t_1,...,x_n/t_n)$ il termine chiuso $ter_2(x_1/t_1,...,x_n/t_n)$.

Un esempio di programma è il seguente che definisce per ricorsione l'operazione di addizione

add[*x*_,0]:=*x*_ ; *add*[*x*_,*s*[*y*_]]:=*s*[*add*[*x*_,*y*_]].

In tale caso i nomi di funzione sono *add* ed *s* , 0 è una costante. Inoltre possiamo interpretare *add* come addizione se accettiamo di muoverci all'interno della terna di Peano (S,s,x_0) il cui dominio è l'insieme *S* dei termini chiusi del tipo 0, *s*[0], *s*[*s*[0]], ... il cui primo elemento x_0 è 0 ed in cui il successore *s* è l'operazione unaria che associa ad ogni termine *t* il termine *s*[*t*]. Se ad esempio si scrive *add*[*s*[0],*s*[*s*[*s*[0]]]] (che noi interpretiamo come 1+3, Mathematica dopo avere effettuato una serie di riscritture fornisce come output *s*[*s*[*s*[*s*[0]]]] che noi interpretiamo come 4.

Si deve notare che Mathematica preferisce non scrivere il segno _ per indicare le variabili nel termine alla destra di una regola e quindi che in Mathematica dobbiamo scrivere tale programma nella forma

add[*x*_,0]:=*x* ; *add*[*x*_,*s*[*y*_]]:=*s*[*add*[*x*,*y*]].

Definizione 3. Un termine chiuso si dice che è scritto in *forma normale* se se non esiste nessuna regola che permetta di riscriverlo. Un *processo di calcolo* finito è una successione finita di termini ognuno dei quali deriva dal precedente tramite una regola di riscrittura e tale che l'ultimo termine è in forma normale. Un *processo di calcolo* infinito è una successione di termini ognuno dei quali deriva dal precedente tramite una regola di riscrittura.

Quando in Mathematica viene fornito un termine chiuso come input, allora l'interprete parte da tale termine e lo riscrive fino a quando è possibile riscriverlo. Il "calcolo", se finisce, restituisce in output l'ultimo termine in forma normale.

Naturalmente il sistema così definito non è deterministico in quanto in ogni passo di una riduzione a forma normale sarebbe possibile applicare più regole di trasformazione. Mathematica utilizza una "strategia" cioè dei criteri con cui scegliere in ogni passo tra le varie possibilità.

Non è compito di questi appunti descrivere in maniera dettagliata quale è la strategia adottata da Mathematica[3]. Comunque, dovendo procedere alla semplificazione-riscrittura di un termine, due regole che vengono applicate in Mathematica sono le seguenti:

---

[3] Il linguaggio Mathematica è in realtà molto più complesso di quanto abbiamo fino ad ora detto. Infatti Mathematica riserva ad alcune parole ed ad alcuni caratteri un significato particolare in accordo con quanto viene fatto dalla comunità dei matematici. Ad esempio la costante *Pi* è utilizzata da Mathematica per denotare il numero π. Inoltre alcune funzioni come *Times*, *Sin*, *Cos*, *Log* sono pre-definite per rappresentare

- si procede alla semplificazione dei sottotermini più interni se questo è possibile
- se esistono più regole di riscrittura per la stesso termine, allora si da la precedenza a quella "più particolare", quando non è chiaro quale sia la regola più particolare si applicano le regole nell'ordine temporale in cui sono state scritte.

Ad esempio, dovendo semplificare il termine Power[Plus[2,3],2], si applica prima la riscrittura a Plus[2,3] ottenendo 5 e poi la riscrittura a Power[5,2] ottenendo 7. Se si scrive il programma
f[x_]:=x^2 /; x>0
f[x_]:=2x /; x<5
allora il termine f[4] verrà riscritto nel termine x^2 e quindi nel termine 4. Si privilegia pertanto la prima regola. Se invece avessi scritto il programma
f[x_]:=2x /; x<5
f[x_]:=x^2 /; x>0
allora f[4] sarebbe stato riscritto in 2 4 e quindi in 8. Naturalmente in entrambi i casi il termine f[-5] si semplifica in -10 in quanto può scattare una sola regola.

10. Un esempio, la derivazione simbolica.

Come esempio mostriamo come potrebbe essere definito un sistema di riscrittura per la derivazione simbolica.

Proposizione 1. Esiste un sistema di riscrittura in Mathematica che permette di definire la funzione derivata (in ambito simbolico)

Dim. Cominciamo con la regola di riscrittura che esprime la regola di derivazione della somma. Potremmo scrivere

der[*f*_+*g*_]:= *der*[*f*]+*der*[*g*]

il cui significato è che se si ha una parola del tipo der[f+g] con f e g parole qualsiasi allora tale parola viene riscritta nella parola der[f]+der[g]. Potremmo aggiungere anche regole del tipo

der[seno[*x*]]:=coseno[*x*], der[coseno[x]]:=-seno[x].

In tale caso se diamo in input la parola
der[seno[x]+coseno[x]]
allora Mathematica la riscrive nella parola
der[seno[x]] +der[coseno[x]]
e poi nella parola
coseno[x]+-seno[x],
infine le regole di riscrittura di libreria trasformano +- in – e quindi si arriva alla parola
coseno[x]-seno[x].
Naturalmente la prima regola scatta ogni volta che viene incontrata una parola che ha al suo interno +. Ad esempio se scrivo der[cane+gatto] ottengo la parola der[cane]+der[gatto]. Se poi aggiungo le regole
der[cane]:=gatto , der[gatto]:=topo.
se scrivo der[cane+gatto] tale espressione verrà riscritta prima in der[cane]+der[gatto] e poi in gatto+topo.
Un sistema di regole più completo per la derivazione simbolica è il seguente dove l'espressione der[f,x] denota la derivata di *f* rispetto ad *x*.

der[f_+g_,x_]:= der[f,x]+der[g,x] ;
der[f_·g_,x_]:=f·der[g,x]+g·der[f,x]

---

l'addizione, le funzioni seno, coseno, logaritmo. In questo caso nella costruzione di un termine esistono vincoli di arità. Ad esempio poiché *Times* è un nome di funzione di due variabili (che denota il prodotto) e *Sin*, *Cos* sono nomi di funzioni ad una variabile, (che denotano le funzioni seno e coseno), allora *Times*[*Sin*[3],*Cos*[Times[2,x_]]] è un termine, che, nella notazione infissa a cui siamo abituati può essere scritto nella forma *Sin*[3]+*Cos*[2+*x*].

der[f_/g_,x_]:=(f·der[g,x]-g·der[f,x])/g^2
der[loga[x_],x_]:=1/x ;
der[seno[x_],x_]:=coseno[x] ;
der[coseno[x_],x_]:=-seno[x]
der[x_^n_,x_]:=n·x^(n-1)
der[x_,x_]:=1 ;
Inoltre è necessaria la seguente regola di riscrittura condizionata
der[*a*_,*x*_]:= 0 /; FreeQ[*a*,*x*]
la quale assicura che la derivata di una espressione *a* rispetto ad *x* è 0 nel caso in cui *x* non compaia come variabile libera in *a* [4]. Si deve comunque osservare che nelle prime tre regole vene ammesso che le variabili f_ e g_ possano essere sostituite con espressioni che normalmente rappresentano funzioni, ad esempio seno[x] e coseno[x]. Questo fatto rientra in quanto detto prima in quanto la *x* che compare in seno[x] e coseno[x] che noi interpretiamo come variabile, in realtà per Matematica è una costante in quanto non termina con il segno _. Non ci soffermiamo ulteriormente su tali questioni e non tentiamo nemmeno di descrivere in modo completo come funziona il sistema di riscrittura di matematica. La cosa importante è solo rendersi conto che Mathematica è basato sulla nozione di sistema di riscrittura.
Concludiamo aggiungendo anche una regola per la derivazione di funzioni composte
der[f_[g_],x_]:=der[g,x]·(der[f[x],x]/.x->g)
dove la stringa /.x->g denota la sostituzione di x con g e quindi der[f[x],x]/.x->g denota il risultato che si ottiene calcolando la derivata di f in x e poi sostituendo al posto della variabile x il valore della funzione g.

Si osservi ancora che le regole di derivazione mettono in evidenza che una definizione f[...]:= B rappresenta una regola di sostituzione che scatta con un processo di "unificazione" e non di "coincidenza". Il processo di unificazione consiste nell'ottenere la coincidenza quando si siano sostituite la variabili con opportuni valori.
Naturalmente Mathematica non è solo un sistema di riscrittura, infatti è dotato anche di:
- un interprete che permette di utilizzare gli usuali strumenti della programmazione procedurale (come i comandi IF THEN ELSE, WHILE, ...)
- una enorme "libreria" di algoritmi numerici, scritti nel linguaggio *C*, che implementano i migliori metodi di calcolo numerico e, più in generale, una enorme libreria che, sotto forma di regole di riscrittura, implementa un pezzo delle conoscenze matematiche (ad esempio i metodi di derivazione e di integrazione simbolica, metodi per la semplificazione di espressioni algebriche ed altro)
- un interfaccia (front-end) che permette l'elaborazione di testi, la grafica, l'animazione, i suoni ed altro.

Naturalmente, poiché il problema dell'integrazione non è decidibile, non è possibile trovare un sistema di riscrittura in matematica relativo all'integrazione simbolica.

Teorema 2. Non può esistere un sistema di regole in Mathematica per l'integrazione simbolica.

[4] FreeQ[*a*,x] è un predicato, pre-definito in Mathematica, che risulta vero se *a* non contiene la lettera x come variabile e falso altrimenti. Ad esempio FreeQ[a[b[x]]a,x] è falso mentre FreeQ[a[b[zxy]]a,x] e FreeQ[x,x] sono vere.

La Biblioteca di Babele
di
Jorge Luis Borges

L'universo (che altri chiama la *Biblioteca*) si compone d'un numero indefinito, e forse infinito, di gallerie esagonali, con vasti pozzi di ventilazione nel mezzo, orlati di basse ringhiere. Da qualsiasi esagono si vedono i piani superiori e inferiori, interminabilmente. La distribuzione degli oggetti nelle gallerie è invariabile. Venticinque vasti scaffali, in ragione di cinque per lato, coprono tutti i lati meno uno; la loro altezza, che è quella stessa di ciascun piano, non supera di molto quella d'una biblioteca normale. Il lato libero dà su un angusto corridoio che porta a un'altra galleria, identica alla prima e a tutte. A destra e a sinistra del corridoio vi sono due gabinetti minuscoli. Uno permette di dormire in piedi; l'altro di soddisfare le necessità fecali. Di qui passa la scala spirale, che s'inabissa e s'innalza nel remoto. Nel corridoio è uno specchio, che fedelmente duplica le apparenze. Gli uomini sogliono inferire da questo specchio che la Biblioteca non è infinita (se realmente fosse tale, perché questa duplicazione illusoria?); io preferisco sognare che queste superfici argentate figurino e promettano l'infinito... La luce procede da frutti sferici che hanno il nome di lampade. Ve ne sono due per esagono, su una traversa. La luce che emettono è insufficiente, incessante.
Come tutti gli uomini della Biblioteca, in gioventù io ho viaggiato; ho peregrinato in cerca di un libro, forse del catalogo dei cataloghi; ora che i miei occhi quasi non possono decifrare ciò che scrivo, mi preparo a morire a poche leghe dall'esagono in cui nacqui. Morto, non mancheranno mani pietose che mi gettino fuori della ringhiera; mia sepoltura sarà l'aria insondabile: il mio corpo affonderà lungamente e si corromperà e dissolverà nel vento generato dalla caduta, che è infinita. Io affermo che la Biblioteca è interminabile. Gli idealisti argomentano che le sale esagonali sono una forma necessaria dello spazio assoluto o, per lo meno, della nostra intuizione dello spazio. Ragionano che è inconcepibile una sala triangolare o pentagonale. (I mistici pretendono di avere, nell'estasi, la rivelazione d'una camera circolare con un gran libro circolare dalla costola continua, che fa il giro completo delle pareti; ma la loro testimonianza è sospetta; le loro parole, oscure. Questo libro ciclico è Dio). Mi basti, per ora, ripetere la sentenza classica: «La Biblioteca è una sfera il cui centro esatto è qualsiasi esagono, e la cui circonferenza è inaccessibile».

A ciascuna parete di ciascun esagono corrispondono cinque scaffali; ciascuno scaffale contiene trentadue libri di formato uniforme; ciascun libro è di quattrocentodieci pagine; ciascuna pagina, di quaranta righe; ciascuna riga, di quaranta lettere di colore nero. Vi sono anche delle lettere sulla costola di ciascun libro; non, però, che indichino o prefigurino ciò che diranno le pagine. So che questa incoerenza, un tempo, parve misteriosa. Prima d'accennare alla soluzione (la cui scoperta, a prescindere dalle sue tragiche proiezioni, è forse il fatto capitale della storia) voglio rammentare alcuni assiomi.

Primo: *La Biblioteca esiste ab aeterno.* Di questa verità, il cui corollario immediato è l'eternità futura del mondo, nessuna mente ragionevole può dubitare. L'uomo, questo imperfetto bibliotecario, può essere opera del caso o di demiurghi malevoli; l'universo, con la sua elegante dotazione di scaffali, di tomi enigmatici, di infaticabili scale per il viaggiatore e di latrine per il bibliotecario seduto, non può essere che l'opera di un dio. Per avvertire la distanza che c'è tra il divino e l'umano, basta paragonare questi rozzi, tremuli simboli che la mia fallibile mano sgorbia sulla copertina d'un libro, con le lettere organiche dell'interno: puntuali, delicate, nerissime, inimitabilmente simmetriche.

Secondo: Il *numero dei simboli ortografici è di venticinque*[5]. Questa constatazione permise, or sono tre secoli, di formulare una teoria generale della Biblioteca e di risolvere soddisfacentemente il problema che nessuna congettura aveva permesso di decifrare: la natura informe e caotica di quasi tutti i libri. Uno di questi, che mio padre vide nell'esagono del circuito quindici novantaquattro, constava delle lettere M C V, perversamente ripetute dalla prima all'ultima riga. Un altro (molto consultato in questa zona) è un mero labirinto di lettere, ma l'ultima pagina dice *Oh tempo le tue piramidi.* È ormai risaputo: per una riga ragionevole, per una notizia corretta, vi sono leghe di insensate cacofonie, di farragini verbali e di

[5] II manoscritto originale non contiene cifre né maiuscole. La punteggiatura è limitata alla virgola e al punto. Questi due segni, lo spazio, e le ventidue lettere d'alfabeto, sono i venticinque simboli sufficienti che enumera lo sconosciuto [N, d. E.].

incoerenze. (So d'una regione barbarica i cui bibliotecari ripudiano la superstiziosa e vana abitudine di cercare un senso nei libri, e la paragonano a quella di cercare un senso nei sogni o nelle linee caotiche della mano... Ammettono che gli inventori della scrittura imitarono i venticinque simboli naturali, ma sostengono che questa applicazione è casuale, e che i libri non significano nulla di per sé. Questa affermazione, lo vedremo, non è del tutto erronea).

Per molto tempo si credette che questi libri impenetrabili corrispondessero a lingue preferite o remote. Ora, è vero che gli uomini piú antichi, i primi bibliotecari, parlavano una lingua molto diversa da quella che noi parliamo oggi; è vero che poche miglia a destra la lingua è già dialettale, e novanta piani piú sopra è incomprensibile. Tutto questo, lo ripeto, è vero, ma quattrocentodieci pagine di inalterabili M C V non possono corrispondere ad alcun idioma, per dialettale o rudimentale che sia. Alcuni insinuarono che ogni lettera poteva influire sulla seguente, e che il valore di M C V nella terza riga della pagina 71 non era lo stesso di quello che la medesima serie poteva avere in altra riga di altra pagina; ma questa vaga tesi non prosperò. Altri pensarono a una crittografia; quest'ipotesi è stata universalmente accettata, ma non nel senso in cui la formularono i suoi inventori.

Cinquecento anni fa, il capo d'un esagono superiore[6] trovò un libro tanto confuso come gli altri, ma in cui v'erano quasi due pagine di scrittura omogenea, verosimilmente leggibile. Mostrò la sua scoperta a un decifratore ambulante, e questi gli disse che erano scritte in portoghese; altri gli assicurò che erano scritte in yiddish. Poté infine stabilirsi, dopo ricerche che durarono quasi un secolo, che si trattava d'un dialetto samoiedo-lituano del guaraní, con inflessioni di arabo classico. Si decifrò anche il contenuto: nozioni di analisi combinatoria, illustrate con esempi di permutazioni a ripetizione illimitata. Questi esempi permisero a un bibliotecario di genio di scoprire la legge fondamentale della Biblioteca.

Questo pensatore osservò che tutti i libri, per diversi che fossero, constavano di elementi uguali: lo spazio, il punto, la virgola, le ventidue lettere dell'alfabeto. Stabilì, inoltre, un fatto che tutti i viaggiatori hanno confermato: non vi sono, nella vasta Biblioteca, due soli libri identici. Da queste premesse incontrovertibili dedusse che la Biblioteca è totale, e che i suoi scaffali registrano tutte le possibili combinazioni dei venticinque simboli ortografici (numero, anche se vastissimo, non infinito) cioè tutto ciò ch'è dato di esprimere, in tutte le lingue. Tutto: la storia minuziosa dell'avvenire, le autobiografie degli arcangeli, il catalogo fedele della Biblioteca, migliaia e migliaia di cataloghi falsi, la dimostrazione della falsità di questi cataloghi, la dimostrazione del catalogo falso, l'evangelo gnostico di Basilide, il commento di questo evangelo, il commento del commento di questo evangelo, il resoconto veridico della tua morte, la traduzione di ogni libro in tutte le lingue, le interpolazioni di ogni libro in tutti i libri.

Quando si proclamò che la Biblioteca comprendeva tutti i libri, la prima impressione fu di straordinaria felicità. Tutti gli uomini si sentirono padroni di un tesoro intatto e segreto. Non v'era problema personale o mondiale la cui eloquente soluzione non esistesse: in un qualche esagono. L'universo era giustificato, l'universo attingeva bruscamente le dimensioni illimitate della speranza. A quel tempo si parlò molto delle Vendicazioni: libri di apologia e di profezia che giustificavano per sempre gli atti di ciascun uomo dell'universo e serbavano arcani prodigiosi per il suo futuro. Migliaia di ambiziosi abbandonarono il dolce esagono natale e si lanciarono su per le scale, spinti dal vano proposito di trovare la propria Vendicazione.

Questi pellegrini s'accapigliavano negli stretti corridoi, profferivano oscure minacce, si strangolavano per le scale divine, scagliavano i libri ingannevoli nei pozzi senza fondo, vi morivano essi stessi, precipitativi dagli uomini di regioni remote. Molti impazzirono. Le Vendicazioni esistono (io ne ho viste due, che si riferiscono a persone da venire, e forse non immaginarie), ma quei ricercatori dimenticavano che la possibilità che un uomo trovi la sua, o qualche perfida variante della sua, è sostanzialmente zero.

Anche si sperò, a quel tempo, nella spiegazione dei misteri fondamentali dell'umanità: l'origine della Biblioteca e del tempo. È verosimile che di questi gravi misteri possa darsi una spiegazione in parole: se il

[6] Prima, per ogni tre esagoni c'era un uomo. Il suicidio e le malattie polmonari hanno distrutto questa proporzione. Fatto indicibilmente malinconico: a volte ho viaggiato molte notti per corridoi e scale levigate senza trovare un solo bibliotecario.

linguaggio dei filosofi non basta, la multiforme Biblioteca avrà prodotto essa stessa l'inaudito idioma necessario, e i vocabolari e la grammatica di questa lingua. Già da quattro secoli gli uomini affaticano gli esagoni... Vi sono cercatori ufficiali, *inquisitori*. Li ho visti nell'esercizio della loro funzione: arrivano sempre scoraggiati; parlano di scale senza un gradino, dove per poco non s'ammazzarono; parlano di scale e di gallerie con il bibliotecario; ogni tanto, prendono il libro piú vicino e lo sfogliano, in cerca di parole infami. Nessuno, visibilmente, s'aspetta di trovare nulla.

Alla speranza smodata, com'è naturale, successe un'eccessiva depressione. La certezza che un qualche scaffale d'un qualche esagono celava libri preziosi, che questi libri preziosi erano inaccessibili, parve quasi intollerabile. Una setta blasfema suggerì che s'interrompessero le ricerche e che tutti gli uomini si dessero a mescolare lettere e simboli, fino a costruire, per un improbabile dono del caso, questi libri canonici. Le autorità si videro obbligate a promulgare ordinanze severe. La setta sparí, ma nella mia fanciullezza ho visto vecchi uomini che lungamente s'occultavano nelle latrine, con dischetti di metallo in un bossolo proibito, e debolmente rimediavano al divino disordine.

Altri, per contro, credettero che l'importante fosse di sbarazzarsi delle opere inutili. Invadevano gli esagoni, esibivano credenziali non sempre false, sfogliavano stizzosamente un volume e condannavano scaffali interi: al loro furore igienico, ascetico, si deve l'insensata distruzione di milioni di libri. Il loro nome è esecrato, ma chi si dispera per i « tesori » che la frenesia di coloro distrusse, trascura due fatti evidenti. Primo: la Biblioteca è così enorme che ogni riduzione d'origine umana risulta infinitesima. Secondo: ogni esemplare è unico, insostituibile, ma (poiché la Biblioteca è totale) restano sempre varie centinaia di migliaia di facsimili imperfetti, cioè di opere che non differiscono che per una lettera o per una virgola. Contrariamente all'opinione generale, credo dunque che le conseguenze delle depredazioni commesse dai Purificatori siano state esagerate a causa dell'orrore che quei fanatici ispirarono. Li sospingeva l'idea delirante di conquistare i libri dell'Esagono Cremisi: libri di formato minore dei normali, onnipotenti, illustrati e magici.

Sappiamo anche di un'altra superstizione di quel tempo: quella dell'Uomo del Libro. In un certo scaffale d'un certo esagono (ragionarono gli uomini) deve esistere un libro che sia la chiave e il compendio perfetto di tutti gli altri: un bibliotecario l'ha letto, ed è simile a un dio. Nel linguaggio di questa zona si conservano alcune tracce del culto di quel funzionario remoto. Molti peregrinarono in cerca di Lui, si spinsero invano nelle piú lontane gallerie. Come localizzare il venerando esagono segreto che l'ospitava? Qualcuno propose un metodo regressivo: per localizzare il libro A, consultare previamente il libro B; per localizzare il libro B, consultare previamente il libro C; e così all'infinito... In avventure come queste ho prodigato e consumato i miei anni.

Non mi sembra inverosimile che in un certo scaffale dell'universo esista un libro totale[7] *; prego gli dèi ignoti che un uomo - uno solo, e sia pure da migliaia d'anni! - l'abbia trovato e l'abbia letto. Se l'onore e la sapienza e la felicità non sono per me, che siano per altri. Che il cielo esista, anche se il mio posto è all'inferno. Ch'io sia oltraggiato e annientato, ma che per un istante, in un essere, la Tua enorme Biblioteca si giustifichi.*

Affermano gli empì che il nonsenso è normale nella Bíblioteca, e che il ragionevole (come anche l'umile e semplice coerenza) vi è una quasi miracolosa eccezione. Parlano (lo so) della «Bíblioteca febbrile, i cui casuali volumi corrono il rischio incessante di mutarsi in altri, e tutto affermano, negano e confondono come una divinità in delirio». Queste parole, che non solo denunciano il disordine, ma lo illustrano, testimoniano generalmente del pessimo gusto e della disperata ignoranza di chi le pronuncia. In realtà, la Biblioteca include tutte le strutture verbali, tutte le variazioni permesse dai venticinque simboli ortografici, ma non un solo nonsenso assoluto. Inutile osservarmi che il miglior volume dei molti esagoni che amministro s'intitola Tuono *pettinato*, un altro Il *crampo di gesso* e un altro Axaxaxas mlö. Queste proposizioni, a prima vista incoerenti, sono indubbiamente suscettibili d'una giustificazione crittografica o allegorica; questa

[7] Ripeto: perché un libro esista, basta che sia possibile. Solo l'impossibile è escluso. Per esempio: nessun libro è anche una scala, sebbene esistano sicuramente dei libri che discutono, che negano, che dimostrano questa possibilità, e altri la cui struttura corrisponde a quella d'una scala.

giustificazione è verbale, e però, *ex hypothesi*, già figura nella Biblioteca. Non posso immaginare alcuna combinazione di caratteri
dhcmrlchtdj
che la divina Biblioteca non abbia previsto, e che in alcuna delle sue lingue segrete non racchiuda un terribile significato. Nessuno può articolare una sillaba che non sia piena di tenerezze e di terrori; che non sia, in uno di quei linguaggi, il nome poderoso di un dio. Parlare è incorrere in tautologie. Questa epistola inutile e verbosa già esiste in uno dei trenta volumi dei cinque scaffali di uno degli innumerabili esagoni - e cosí pure la sua confutazione. (Un numero n di lingue possibili usa lo stesso vocabolario; in alcune, il simbolo *biblioteca* ammette la definizione corretta del *sistema duraturo e ubiquitario di gallerie esagonali,* ma biblioteca sta qui per *pane, o* per *piramide, o* per qualsiasi altra cosa, e per altre cose stanno le sette parole che la definiscono. Tu, che mi leggi, sei sicuro d'intendere la mia lingua?)

Lo scrivere metodico mi distrae dalla presente condizione degli uomini, cui la certezza di ciò, che tutto sta scritto, annienta o istupidisce. So di distretti in cui i giovani si prosternano dinanzi ai libri e ne baciano con barbarie le pagine, ma non sanno decifrare una sola lettera. Le epidemie, le discordie eretiche, le peregrinazioni che inevitabilmente degenerano in banditismo, hanno decimato la popolazione. Credo di aver già accennato ai suicidi, ogni anno piú frequenti. M'inganneranno, forse, la vecchiezza e il timore, ma sospetto che la specie umana - l'unica - stia per estinguersi, e che la Biblioteca perdurerà: illuminata, solitaria, infinita, perfettamente immobile, armata di volumi preziosi, inutile, incorruttibile, segreta.

Aggiungo: *infinita.* Non introduco quest'aggettivo per un'abitudine retorica; dico che non è illogico pensare che il mondo sia infinito. Chi lo giudica limitato, suppone che in qualche luogo remoto i corridoi e le scale e gli esagoni possano inconcepibilmente cessare; ciò che è assurdo. Chi lo immagina senza limiti, dimentica che è limitato il numero possibile dei libri. Io m'arrischio a insinuare questa soluzione: *La Biblioteca è illimitata e periodica.* Se un eterno viaggiatore la traversasse in una direzione qualsiasi, constaterebbe alla fine dei secoli che gli stessi volumi si ripetono nello stesso disordine (che, ripetuto, sarebbe un ordine: l'Ordine). Questa elegante speranza rallegra la mia solitudine[8] .
1.Forse questa opinione è eretica. San Tommaso d'Aquino (citato da Bertrand Russeli in *A Historiy of Western Philosophy,* New York, Simon and Schuster, 1945, p. 458) afferma che Dio non può far sì che un uomo non abbia l'anima. Ma questa può non essere una limitazione reale della. Sua potenza, bensì solo una conseguenza del fatto che l'anima dell'uomo è immortale e quindi indistruttibile.
1941, Mar della Plata.

[8] I Letizia Alvarez de Toledo ha osservato che la vasta Biblioteca è inutile; a rigore, basterebbe un solo *volume,* di formato comune, stampato in corpo nove o in corpo dieci, e composto d'un numero infinito di fogli infinitamente sottili. (Cavalieri, al principio del secolo xvii, affermò che ogni corpo solido è la sovrapposizione d'un numero infinito di piani). Il maneggio di questo serico *vademecum* non sarebbe comodo: ogni foglio apparente si sdoppierebbe in altri simili; l'inconcepibile foglio centrale non avrebbe rovescio.