PER ISTRUIRE UN ROBOT

ovvero, come costruirsi una logica

Giovanni Sambin
sambin@math.unipd.it

Capitolo 3 - Gli usi della logica

In questo terzo capitolo vogliamo mettere a frutto quello che abbiamo imparato sulla logica. Quello che abbiamo imparato pensando di istruire il robot ora vedremo ci sarà utile per descrivere, ad esempio nella vita di tutti i giorni, cose come la struttura delle parentele, o, in campo matematico, le strutture algebriche. Arriveremo così a parlare di metodo assiomatico.

1. Esempio: la struttura delle parentele.

Cominciamo a vedere come si usa la logica nei casi specifici. Applicheremo ciò che abbiamo imparato fino adesso, e cioè il concetto di segno, espressione, le deduzioni, terzo escluso, negazione, etc. ad un esempio. Questo esempio deve essere noto a voi dal punto di vista di conoscenza in modo da non avere dubbi sulla materia e però ignoto quanto alla sua scrittura in termini formali. Quello che vogliamo fare è, in altri termini, usare la logica sia come linguaggio –nel senso simbolico e cioè con i connettivi e i quantificatori– sia come deduzione per descrivere un campo che noi tutti conosciamo: la struttura dei gradi di parentela tra esseri umani e l'uso di concetti e parole come figlio, padre, madre, zio, eccetera. Scelgo qualcosa che tutti conosciamo benissimo, in modo che si possa concentrare l'attenzione sul modo di metter a frutto la logica.

Nello stesso tempo, darò anche alcune nuove nozioni sulla logica, che vedremo sì nel caso specifico, ma che chiaramente si potranno applicare in generale. Ad esempio, vedremo le regole di deduzione per trattare il predicato di uguaglianza.

Immaginate di essere già degli esperti e di dover costruire una base di dati per il vostro comune di residenza che racconti l'anagrafe. Per la base di dati si deve specificare un linguaggio con cui raccontare certe informazioni. Naturalmente, intendo non una lista di dati, ma una base di dati intelligente, e cioè che possa rispondere a domande del tipo "chi sono i cugini di ..." oppure "quanti figli ha ..." e che riesca anche a far deduzioni su queste cose.

Cosa ci serve per formalizzare o assiomatizzare un campo di conoscenza, e cioè "tradurlo" in proposizioni, calcoli, assiomi, etc. ?

Bisogna innanzitutto specificare un dominio; nel nostro caso sarà costituito o da tutti gli esseri umani, o dai residenti del nostro comune. In ogni caso, supponiamo che sia un insieme, o in altre parole, supponiamo di sapere che senso ha scrivere $\forall x \in D \ldots$ D'ora in poi utilizzeremo le variabili $x, y, z, t, \ldots$ per indicare persone, individui, senza specificare il dominio.

In altre parole non scriveremo più $\forall x \in D$ in quanto superfluo trattandosi sempre della stessa cosa.

Il linguaggio della logica automaticamente prevede inoltre l'uso dei connettivi $\&, \vee, \rightarrow$ e dei quantificatori $\forall, \exists$. Questi vanno applicati a proposizioni o funzioni proposizionali; nello studio teorico come abbiamo fatto finora, non sono mai state specificate. Ora questo va affrontato, ora diventa necessario specificare di quali proposizioni e quali relazioni vogliamo parlare.

Per relazione intendo semplicemente una funzione proposizionale; se tale funzione proposizionale ha un solo argomento, cioè una sola variabile libera, allora spesso si chiama un sottoinsieme, o un predicato. Il nome relazione spesso si applica a funzioni proposizionali con due argomenti, ma nulla ci impedisce di considerare anche relazioni con più argomenti.

Quali sono le proposizioni, le funzioni proposizionali (alias relazioni) da cui ci conviene partire? Notate che ho detto "ci conviene partire" nel senso che è una scelta nostra. Siamo noi che scegliamo come organizzare la conoscenza in modo che ci venga più utile. Per parlare della struttura delle parentele dobbiamo scegliere quali saranni i *concetti primitivi* e quali saranno i *concetti definiti*, cioe' quelli ottenibili a partire dai primi. Inoltre, per i concetti primitivi dobbiamo decidere quali sono le proposizioni che vogliamo prendere come assunzioni primitive, cioè assiomi, e quali proprietà possiamo invece dimostrare.

Quindi da quale idea dobbiamo partire per descrivere, mediante essa e nel modo più semplice, tutte le relazioni di parentela, in tutti i loro gradi? Un'idea potrebbe essere partire da "essere figlio di ..." che tradotto diventa x è figlio di y e di z (dopo una discussione in aula, si è visto che conviene specificare sia la madre che il padre). Dico tradotto perché nel linguaggio italiano non si usano variabili in modo esplicito, però sappiamo benissimo il significato di "essere figlio di ... e di ...", ad esempio capiamo benissimo cosa vuol dire la proposizione "Fabio è figlio di Maria e di Antonio". Nel linguaggio della logica al posto degli individui specifici Fabio, Maria e Antonio ci possono essere delle variabili, per cui il predicato "essere figlio di ... e di ..." si esprime anche con x è figlio di y e di z. A questo punto, si può usare una notazione più breve, e scrivere semplicemente $F(x, y, z)$ per indicare che x è figlio di y e di z. Quindi:

$$F(x, y, z) =_{def} x \text{ è figlio di } y \text{ e di } z$$

Da questa dobbiamo ora essere in grado di descrivere tutte le altre parentele. Per esempio, possiamo dire che ogni x ha due genitori, che i genitori non sono la stessa persona, e quale tra i due è la femmina o il maschio? Per affermare che ogni individuo x ha due genitori basta dire che esistono due individui tali che x è loro figlio. Quindi basta scrivere

$$\forall x \, \exists y \, \exists z \, F(x, y, z)$$

Questo però non esclude che y e z siano la stessa persona. Per indicare quale tra i due è il maschio e quale la femmina potremmo sfruttare la posizione, cioè metteremo la femmina prima del maschio. Quindi in $F(x, y, z)$, y è la femmina mentre z è il maschio.

Ma si vede subito che questo non basta per distinguere i sessi, nel senso che in tal modo riusciamo a dire solo il sesso degli individui che hanno almeno un figlio, e come sappiamo bene non tutti gli individui hanno dei figli. In altri termini, partendo dal predicato "essere figlio di ..." non possiamo ottenere una caratterizzazione del predicato "essere maschio" ed "essere femmina".

Si potrebbe pensare ad una funzione che per ogni x dice se è maschio o se è femmina. Forse il modo più semplice sarebbe quello di aggiungere due predicati, uno $U(x)$ che dice "x è un maschio, o uomo", e uno $D(x)$ che dice "x è una femmina, o donna". Invece proviamo una nuova via: introduciamo una funzione che ad x associa $s(x)$ (cioè il sesso di x) e che ovviamente può assumere due soli valori; inoltre è definito su tutti gli individui ed è una vera funzione perchè ognuno di noi ha un solo sesso. Di solito si pensa a "maschio" e "femmina" come proprietà o come predicati. Ma si può passare dall'uno all'altro modo di ragionare; vediamo come. Supponiamo di avere la funzione

$$x \longmapsto s(x) = \begin{cases} m & \text{se } x \text{ è maschio} \\ f & \text{se } x \text{ è femmina} \end{cases}$$

Si noti che "x è maschio" e "x è femmina" fanno parte del metalinguaggio perché dentro al sistema formale abbiamo solo la funzione s. Si potrebbe introdurre solo $D(x)$ e poi definire "essere maschio" come $\neg D(x)$, cioè per dire "x è un uomo" e indicarlo con $U(x)$ non devo necessariamente introdurlo come predicato primitivo, ma posso ottenerlo da $D(x)$ usando il connettivo $\neg$. Mi sembra che comunque convenga introdurli entrambi, per rendere l'espressione formale più vicina a quel che si fa nella lingua d'ogni giorno.

Come si passa dalla funzione al predicato? Data la funzione s, poniamo

$$\begin{array}{lll} U(x) \equiv s(x) = m & & x \text{ è maschio} \\ D(x) \equiv s(x) = f & & x \text{ è femmina} \end{array}$$

Questo è esatto perché ad esempio $s(x) = m$ è una proprietà ed x la soddisfa quando è un maschio. Essendo s una funzione e supponendo $m \neq f$ ne segue che

$$(\forall x)\,((U(x) \lor D(x))\,\&\,\neg(D(x)\,\&\,U(x))$$

ovvero ciascuno o è maschio o è femmina e non può essere le due cose contemporaneamente.

Viceversa, come si passa dai predicati alla funzione? Quello che vogliamo trovare è questo: dati i predicati $U(x)$ e $D(x)$, trovare una funzione s definita su tutti gli x e che soddisfi:

$$\begin{array}{l} U(x) \vdash s(x) = m, \text{ cioè: } x \text{ uomo comporta il sesso di } x \text{ è } m \\ D(x) \vdash s(x) = f \end{array}$$

Quindi dato x, vogliamo spiegare come calcolare $s(x)$, ovvero ottenere se deve essere uguale a m o a f. Il metodo è questo: si usa l'assioma

$$(\forall x)((U(x) \lor D(x))\,\&\,\neg(D(x)\,\&\,U(x)))$$

che vale per ogni x e, dunque, anche ad esempio per il nostro x specifico di cui vogliamo calcolare il sesso; sia ad esempio x_0. Vale

$$(U(x_0) \lor D(x_0)\,\&\,\neg(D(x_0)\,\&\,U(x_0))$$

Dunque, se è vera $U(x_0)$, la funzione varrà m; se è vera $D(x_0)$ la funzione varrà f. Per poter avere la funzione, abbiamo bisogno di decidere in modo algoritmico quale tra $U(x_0)$ e $D(x_0)$ è vera.

A questo punto, grazie alle assunzioni fatte, posso finalmente esprimere il fatto che ogni madre y è una donna e ogni padre z è un maschio. Se scelgo

$$\forall y\ (\exists w \exists t\, F(w, y, t) \rightarrow\ D(y))$$

come assioma, mi dice che "per ogni individuo y, se esistono due individui w e t tali che w è figlio di y e di t, allora y è una donna". Notate che con ciò abbiamo anche espresso il fatto che la y, cioè il secondo argomento in $F(x, y, z)$, è la madre. Analogamente per il padre, che deve essere maschio

$$\forall z\ (\exists x \exists w\, F(x, w, z) \rightarrow\ U(z))$$

Non so se l'avete notato, ma così facendo ho anche specificato che se $F(x, y, z)$ allora $y \neq z$; o meglio, questo fatto si può dimostrare a partire dalle assunzioni fatte. Infatti da $F(x, y, z)$ segue subito che $\exists x\, \exists z\, F(x, y, z)$, e quindi per l'assunzione segue anche subito che $D(y)$. Nello stesso modo, da $F(x, y, z)$ segue subito che $\exists x\, \exists y\, F(x, y, z)$ e quindi $U(z)$. A questo punto, se fosse $y = x$ avrei anche $D(y)\&U(y)$ (vedi in seguito le regole sull'uguaglianza) il che è assurdo in quanto va contro l'assunzione $\neg\exists x(U(x)\&D(x))$. Quello che voglio dire è che non serve assumere $y \neq z$ perché segue dalle assunzioni già fatte, cioè sono in grado di dimostrare che $\forall x, y, z\, (F(x, y, z) \rightarrow\ y \neq z)$.

Ma quando dico "dimostrare" $y \neq z$ a partire da $F(x, y, z)$, qui intendo proprio una dimostrazione nel calcolo dei sequenti LJ o LK, del sequente $\Gamma, F(x, y, z) \vdash\ y \neq z$, dove Γ può contenere tutti gli assiomi che abbiamo adottato finora. Lo lascio come esercizio (di abilità e di pazienza, nel saper organizzare sia l'argomentazione in generale, sia i suoi dettagli formali). Quindi il lavoro fatto sino ad ora è in grado di dimostrare questo o meglio istruendo il nostro computer con le regole e gli assiomi che abbiamo appena introdotto esso è in grado di dirci che se y e z sono madre e padre di qualche individuo, allora $y \neq z$.

Un consiglio: si prenda un grande foglio di carta, e si scriva *tutta* la derivazione nei minimi dettagli, per rendersi conto che davvero si usano solo le regole formali. A quel punto si vedrà bene che serve un passaggio

$$U(y), y = z \vdash U(z)$$

Si tratta di una assunzione non ancora introdotta e spiegata, sulla uguaglianza. Quindi diamo per assunto che ci sia nel nostro dominio una funzione proposizionale, per la quale usiamo il segno =, che lega due individui quando sono uguali, che per esempio nel caso degli esseri umani corrisponde all'identità: dichiariamo che due esseri umani sono uguali quando sono lo stesso individuo.

In generale, ogni volta che si ha bisogno del concetto di uguaglianza (cioè, quasi sempre, e certamente noi ora con le parentele), si assume che la relazione di *uguaglianza* soddisfi le seguenti:

- riflessiva
$$\vdash x = x$$
- simmetrica
$$\Gamma, y = x \vdash x = y$$
- transitiva
$$\Gamma, x = y, y = z \vdash x = z$$

- sostituzione in proposizioni

$$\Gamma, x = y, A(x) \vdash A(y)$$

  e lo stesso nel caso di più variabili;
- sostituzione in termini (vedi def. nel cap. 4)

$$\Gamma, x = y \vdash f(x) = f(y)$$

  e lo stesso nel caso di più variabili.

Quando si scrive che $\vdash x = x$, è sottointeso che questo valga a partire da zero assunzioni, e quindi si può applicare la regola di $\forall$-formazione per ottenere $\vdash \forall x\,(x = x)$. Lascio come esercizio di derivare similmente che valgono $\vdash \forall x, y\,(x = y \to y = x)$ e $\vdash \forall x, y, z(x = y \,\&\, y = z \to x = z)$.

Riprendiamo il nostro esercizio e definiamo –partendo sempre dalle assunzioni fatte– il concetto di padre e madre. Vediamo per prima cosa come posso dire "x è figlio di w" senza specificare se w è padre o madre, e sempre a partire dalla sola relazione $F(x, y, z)$. Per dire che x è figlio di w, basta dire w è o il padre o la madre di x, e questo equivale a dire che x è figlio di un certo y e di un certo z, e che vale che o $w = y$ o che $w = z$. Quindi posso definire "x è figlio di w" come

$$F'(x, w) =_{def} \exists y \, \exists z \; (F(x, y, z)\&(w = y \;\vee\; w = z))$$

In questa proposizione le variabili libere –ovvero non quantificate– sono solo x e w, quindi F' è una relazione che lega x a w.

Lo stesso concetto "padre di x" si può esprimere in due modi: "essere padre di x" oppure "l'individuo che è padre di x". Le due cose sono diverse in quanto la prima è una proprietà, mentre la seconda è una persona, un individuo. Meglio: "essere padre di x" è una funzione proposizionale, possiamo indicarla con $P(x, y)$, mentre l'altra è un certo individuo. Risultano ovviamente legate tra di esse perché l'estensione della proprietà "essere padre di x" è un sottoinsieme formato da un solo individuo, cioè appunto l'unico y per cui "y è padre di x". Cioè noi sappiamo che "essere padre di x" è una funzione proposizionale soddisfatta da un solo individuo e quindi spesso confondiamo la proprietà con l'unico individuo che la soddisfa.

Piccola osservazione: il fatto che in $P(x, y)$ si usi la variabile y, la stessa che si usava in $F(x, y, z)$ per indicare la madre, e quindi femmina mentre il maschio era z, non costituisce alcun problema, né dovrebbe essere fonte di confusione. Si deve stare attenti a non confondersi, perché si può sì anche metter z, ma non si sbaglia se si mette y, perché si sta descrivendo una proprietà che non dipende da quella indicata con $F(x, y, z)$. Si deve capire, al di là dei segni che si usano, cosa si sta dicendo; che si usi y o z oppure un asterisco, l'importante è capire che quella è una variabile, vincolata o libera, il che è molto più importante del nome che uso.

Ritorniamo al nostro esercizio; in simboli, la distinzione si può fare così: $P(x, y)$, che indica la funzione proposizionale "y è padre di x, l'associo all'insieme degli y tali che soddisfano la proprietà e che indico con $\{y \in D : P(x, y)\}$. Per ciascun x, questo è il sottoinsieme dei padri di x; sappiamo bene che questo sottoinsieme è formato da un solo individuo, ma è comunque un sottoinsieme, cioè l'estensione di una proprietà. Invece, il padre di x, inteso come individuo, può essere dato da una funzione che ad ogni individuo x associa quel ben determinato individuo che è il padre di x, e che possiamo indicare con $p(x)$.

Si noti che vi è un legame stretto tra $p(x)$, "il padre di x", e $P(x, y)$, "essere padre di x", ma che non sono la stessa cosa. La prima scrittura indica un individuo del nostro dominio in funzione di x, mentre la seconda è una proposizione su x, e quindi non possono essere uguali. Il padre di x è un elemento, mentre gli y che soddisfano $P(x, y)$ sono un sottoinsieme, anche se sappiamo che questo sottoinsieme ha come unico elemento $p(x)$. La differenza quindi è circa la stessa che c'è tra un elemento d e il sottoinsieme $\{d\}$ che ha d come unico elemento.

Faccio un esempio per chiarire meglio il concetto. Il padre di Maria Antonietta è un solo individuo che si chiama Salvatore. La proprietà "essere padre di Maria Antonietta" la capiamo anche senza sapere chi è il padre. Come tutte le proprietà essa determina un sottoinsieme formato dagli individui che la soddisfano, però nel nostro caso questa proprietà è soddisfatta dal solo individuo che corrisponde al nome di Salvatore, quindi tutto ciò, in notazione matematica, è il singoletto {Salvatore}. Se invece della proprietà "essere padre" avessi considerato la proprietà "essere zio" non avrei più potuto fare questa identificazione, perché in generale un individuo ha più di uno zio (quindi non un singoletto), oppure anche zero zii (il sottoinsieme vuoto).

Chiusa questa parentesi vediamo come si può dire "essere zio di x". Cominciamo dapprima a definire l'essere padre, utilizzando il simbolo $P(x, y)$ per indicare "y è padre di x". Vado a definirlo nel seguente modo:

$$P(x, y) =_{def} \exists z\, F(x, z, y)$$

cioè "y è padre di x se esiste un individuo z tale che x è figlio di z e di y".

Notate come in questa proprietà le variabili x e y compaiono libere e come non sia importante il nome che attribuisco ad esse ma la posizione che occupano. La variabile z posso cambiarla come voglio, cosa che invece non posso fare con y e con x essendo variabili libere a meno che non li cambi sia in $P(x, y)$ che in $F(x, y, z)$. In altre parole la x e la y che compaiono a sinistra devono essere "uguali" a quelle che compaiono a destra. Similmente si definisce la madre.

Il punto più importante è che, nel caso di madre e padre, trattandosi di un solo individuo conviene trovare un modo più veloce per indicarlo e specificare che è unico. In altri termini devo poter dire "il padre di x è unico". Si noti che qui stiamo cercando di scrivere (con la logica che abbiamo studiato finora) una cosa che di solito si dice a parole.

L'assioma è: "ogni figlio ha un solo padre", che ci è molto famigliare; ora, però, dobbiamo dirlo al robot. I mezzi per istruirlo li abbiamo, però non possiamo dirglielo come lo diciamo a livello di metalinguaggio; glielo diremo in un modo equivalente, più adatto al robot.

Come facciamo? Come facciamo a dire che ogni x ha un solo padre? Bisognerebbe saper scrivere "esiste un unico y" e il modo c'è: $\forall x \exists! y P(x, y)$, che si legge "per ogni x, esiste un unico y con y padre di x", solo che $\exists!$ non l'abbiamo mai definito. Come fare allora? Se abbiamo anche la relazione di uguaglianza, oltre a dire $P(x, y)$, possiamo anche dire che ogni z che sia padre di x è in realtà uguale a y. Cioè poniamo:

$$\exists! y P(x,y) \equiv \exists y(P(x,y) \ \& \ \forall z(P(x,z) \rightarrow y = z))$$

e questo per definizione vuol dire "esiste un solo y tale che $P(x,y)$". Quindi "ogni individuo ha un solo padre" si scrive

$$\forall x \exists! y P(x,y)$$

Si notino due cose: questa scrittura dice che tutti gli z sono uguali a y, ma questo y non è un y generico; è l'y che dichiariamo esistere. L'azione di questo quantificatore $\exists y$ è su tutta la formula $P(x,y) \ \& \ \forall z(P(x,z) \rightarrow y = z)$. Quindi quando diciamo "ogni z è uguale a y" è quell'y lì che dichiariamo esistere: esiste un y a cui sono uguali tutti gli z padri di x. (Si tenga presente che potevano esserci altri modi equivalenti per dirlo). Questo, dunque, sarà un assioma. È chiaro che d'ora in poi abbrevieremo $\exists! y$, "esiste un unico" in questo modo, e non occorrerà ridefinirlo perché ormai sappiamo che cos'è.

Poiché la funzione proposizionale $P(x,y)$ –così come $M(x,y)$– è soddisfatta per ogni dato x da un solo individuo, posso indicarla più brevemente con $p(x)$ –e rispettivamente con $m(x)$– cioè $p(x)$ è l'unico individuo che soddisfa la proprietà $P(x,y)$ ovvero $p(x) = y$. La p e la m sono funzioni, nel senso che quando sono applicate ad un individuo x, danno un individuo che si indica con $p(x)$ (e $m(x)$).

Poi potremmo dire che padre e madre non possono coincidere. Abbiamo scritto "ogni individuo ha un solo padre" e "ogni individuo ha una sola madre". Come possiamo dire che "padre e madre di y non coincidono"?

Potremmo dire che l'insieme delle madri intersecato all'insieme dei padri è vuoto: (insieme delle madri) $\cap$ (insieme dei padri) $= \emptyset$. Qual è intuitivamente l'insieme delle madri? Per dire che y è una madre, basta dire che esiste un x di cui y è madre, cioè $\exists x M(x,y)$. Qui y è una variabile non quantificata, e quindi dire "y sta nell'insieme delle madri" è esattamente come dire $\exists x M(x,y)$. Quindi non c'è bisogno di introdurre nuovi insiemi, come l'insieme delle madri. Dunque "l'insieme delle madri intersecato all'insieme dei padri è vuoto" si può dire così:

$$\forall x \forall y(P(x,y) \rightarrow \neg \exists z M(z,y))$$

Leggiamo quel che abbiamo scritto: qualunque sia x e qualunque sia y, se y è padre di x allora non esiste uno z tale che y è la madre di z. Oppure possiamo dire così:

$$\forall y(\exists x P(x,y) \ \& \ \exists z M(z,y) \rightarrow \bot)$$

Perché implica falso? Perché è un modo per dire che non è possibile: qualunque sia y non è possibile che sia padre di un x e madre di uno z - si noti che qui x e z sono quantificati: $\exists x \ \exists z$.

Esercizio: vedere che legame c'è tra

$$\forall y(P(x,y) \rightarrow \neg \exists z M(z,y))$$

e

$$\exists x P(x,y) \ \& \ \exists z M(z,y) \rightarrow \bot$$

Si noterà che sono molto simili.

Supponiamo di aver dimostrato che

$$\forall x(Ax \to B) \; sse \; \exists x Ax \to B$$

a patto che x non compaia libera in B. Allora in $\neg\exists z M(z,y)$ y non compare libera, quindi $\forall x$ lo possiamo portare davanti e diventa un $\exists x$:

$$\exists x P(x,y) \to \neg\exists z M(z,y))$$

Ora vediamo che sono proprio equivalenti, perché $A \to \neg B$ equivale a $A\&B \to \bot$ (esercizio; è un caso particolare di $A \to (B \to C) \doteq A\&B \to C$). Dimostriamo ora che $\forall x(Ax \to B) \doteq \exists x Ax \to B$ (esercizio; qui è importantissimo che x non compaia in B altrimenti non potremmo usare $\exists$-formazione e $\forall$-formazione).

Potremmo definire anche il concetto di zio, nonno e cugino ma per far ciò occorre definire prima il concetto di fratello e sorella. La cosa è alquanto semplice; due individui sono fratelli –o fratello e sorella o ancora sorelle– se innanzi tutto non sono la stessa persona e se hanno i genitori in comune. Quindi x e y sono fratelli se w e t sono genitori di x e genitori di y contemporaneamente ed inoltre se $y \neq x$. Poiché noi come concetto primitivo abbiamo scelto l'essere figlio, dobbiamo obbligatoriamente definire il concetto di "essere fratelli" in termini di "essere figlio". Quindi x e y sono fratelli se esistono due individui w e t tali che x e y sono figli di w e di t e $y \neq x$. Utilizzando il simbolo $FS(x,y)$ per dire "y è fratello (sorella) di x" vado a definire la proprietà "essere fratello (sorella) di x" nel seguente modo:

$$FS(x,y) =_{def} \exists w \exists t \; (F(x,w,t)\&F(y,w,t)) \,\&\, x \neq y$$

oppure anche

$$FS(x,y) =_{def} p(x) = p(y) \wedge m(x) = m(y)\&\, x \neq y$$

Possiamo allora definire "x è fratello di y" come $FS(x,y) \wedge U(x)$ e similmente 'x è sorella di y" come $FS(x,y) \wedge D(x)$.

Adesso siamo in grado di definire il predicato "essere zio". Un individuo y è lo zio di x se y è fratello (sorella) di uno dei genitori di x. Per dire "y è zio di x" uso il simbolo $Z(x,y)$ e vado a definirlo nel seguente modo:

$$Z(x,y) =_{def} \exists w \; (FS(w,y)\&\, (P(x,w) \vee M(x,w)))$$

che leggo *"y è zio/a di x se esiste un individuo w tale che y è fratello (sorella) di w e w è padre di x oppure w è madre di x"*. Sfruttando la notazione "diretta" $p(x)$, $m(x)$ posso definirlo più semplicemente come

$$Z(x,y) =_{def} FS(p(x),y) \vee FS(m(x),y)$$

cioè *"y è zio/a di x se y è fratello (sorella) di $p(x)$ o se y è fratello (sorella) di $m(x)$"*.

Posso definirlo ancora in un terzo modo; introduco il concetto di genitore

$$G(x,y) =_{def} P(x,y) \vee M(x,y)$$

che leggo *"y è genitore di x se y è padre di x oppure se y è madre di x"*, e mediante esso vado a definire il concetto di zio in un terzo modo

$$Z(x,y) =_{def} \exists w \, (G(x,w)\&FS(w,y))$$

che leggo *"y è zio/a di x se esiste un individuo w tale che w è genitore di x e y è fratello (sorella) di w"*.

Questi tre modi per definire il concetto di zio sono tutti corretti. Tocca a me decidere quale scegliere in base alla pesantezza. Solitamente si sceglie quello contenente meno notazioni.

Passiamo ora al concetto di nonno/a. Il nonno/a di x è l'individuo y tale che y è padre (madre) del padre di x o della madre di x. Utilizzo il simbolo $N(x,y)$ per dire "y è nonno/a di x" e vado a definirlo di seguito

$$N(x,y) =_{def} \; y = p(p(x)) \vee y = p(m(x)) \vee y = m(p(x)) \vee y = m(m(x))$$

che leggo *"y è nonno/a di x se y è il padre di $p(x)$ o se y è il padre di $m(x)$ o se y è la madre di $p(x)$ o se y è la madre di $m(x)$"*.

Osservate com'è pesante questa definizione nel senso che è troppo lunga. Se utilizzassi il concetto di genitore, invece che quello di padre e madre, avrei

$$N(x,y) =_{def} \; \exists w\,(G(x,w) \& G(w,y))$$

che leggo *"y è nonno/a di x se esiste un individuo w tale che w è genitore di x e y è genitore di w"* che ovviamente risulta essere molto più semplice della precedente.

Più in generale potremmo definire il concetto di ascendente di grado n-esimo. Potremmo utilizzare qualcosa di ricorsivo, per esempio:

$$\begin{aligned} A^n(x,y) \; &\equiv \; \exists w\,(M(w,x) \; \vee \; P(w,x) \,\&\, A^{n-1}(w,y) \\ &\equiv \; x \text{è ascendente } n\text{-esimo di } y. \end{aligned}$$

con assioma: $A^0(x,x)$ cioè x è ascendente 0-esimo di x.
Potrebbe essere corretto. Stiamo definendo per induzione "essere ascendente di grado n":

1) definiamo il grado zero:

$$A^0(x,y) \; \longleftrightarrow \; x = y$$

cioè: "sono io ascendente di me stesso".

2) poi il passo 1 sarà:

$$A^1(x,y) \; \equiv \; \exists w((M(w,x) \; \vee \; P(w,x)) \,\&\, A^0(w,y))$$

Dunque da 1) segue che $w = y$ e quindi $\exists w((M(w,x) \; \vee \; P(w,x)) \,\&\, w = y)$ e otteniamo $M(y,x) \; \vee \; P(y,x)$. Questo passaggio formale appena descritto non l'abbiamo ancora trattato, ma è un caso particolare dell'equivalenza:

$$(\exists z)\,(z = y \,\&\, A(z)) \; \doteq \; A(y)$$

la cui derivazione formale lasciamo come esercizio.

Quindi si può sviluppare questa teoria di ascendente n-esimo ed otteniamo come casi particolari:
$A^1(x,y)$"x genitore di y"
$A^2(x,y)$ "x nonno di y"
$A^3(x,y)$ "x bisnonno di y "

Andiamo al concetto di cugino. Sappiamo che i cugini di x sono i figli degli zii di x quindi y è cugino di x se y è figlio di w e w è zio di x. Utilizzo il simbolo $Cg(x,y)$ per dire "y è cugino di x" e lo vado a definire di seguito

$$Cg(x,y) =_{def} \; \exists w\,(Z(x,w) \& G(y,w))$$

che leggo *"y è cugino di x se esiste un individuo w tale che w è zio di x e w è genitore di y"*.

Un particolare alquanto curioso: se nella definizione di cugino andiamo a sostituire le definizioni di zio e genitore otteniamo come risultato che due individui sono cugini se hanno due nonni in comune (alcuni studenti se ne rendono conto qui per la prima volta!) . Quindi "x è cugino di y" si scrive anche:

$$Cg(x,y) \equiv (\exists z)\,(A^2(z,x)\,\&\,A^2(z,y))\,\&\,\neg\,FS(x,y)$$

Riassumendo, abbiamo visto qualche esempio che illustra ciascuna delle seguenti osservazioni:

1) si possono dare delle definizioni;
2) concetti equivalenti corrispondono ad equivalenze formali;
3) si possono usare funzioni per costruire o indicare individui;
4) si scelgono degli assiomi che si possono modificare man mano che si acquisisce conoscenza del campo;
5) gli stessi assiomi possono avere più di una interpretazione.

Noi stiamo, in un certo senso, assiomatizzando la struttura delle parentele, ma alla fine quello che descriviamo è un albero (o meglio delle foreste) con tanti individui, ciascuno dei quali ha due ascendenti immediati e un sesso (potremmo ad esempio usare un colore rosso o verde). *Tutte le definizioni sono solo nomi con cui indichiamo legami tra punti della "nostra foresta".* Qualunque struttura fatta così soddisfa gli assiomi che abbiamo scritto.

Questo è un vantaggio e allo stesso tempo una debolezza della scrittura in simboli: una debolezza perché non dice tutto quello che noi avevamo in mente; un vantaggio perché è chiaro che cosa abbiamo dato come istruzione al robot. Quindi quel che il robot sa è passibile di dimostrazione matematica, "di conoscenza con assoluta certezza". (Ma ad esempio con questi assiomi, il robot non è in grado di definire "marito" e "moglie").

2. Alcuni esempi dalla matematica.

2.1. Insiemi ordinati. Conclusi gli esempi sulle parentele, vediamo ora come le stesse cose che abbiamo visto, le stesse osservazioni, gli stessi analoghi esempi si ritrovano in matematica e in che senso. Come facciamo per esempio a descrivere con il linguaggio formale della logica un *insieme ordinato*? Cominiciamo con un esempio specifico: vogliamo descrivere $(\mathbf{N}, \leq)$, la struttura dei numeri naturali con il loro ordine usuale. Allora diremo che c'è una relazione $x \leq y$ che inoltre soddisfa un po' di proprietà. Per cominciare, dobbiamo dire che c'è un ordine parziale, quindi dobbiamo specificare che

(1) proprietà riflessiva $\quad \vdash x \leq x$
(2) proprietà transitiva $\quad x \leq y, y \leq z \vdash x \leq z$
(3) proprietà antisimmetrica $\quad x \leq y, y \leq x \vdash x = y$

Poi dobbiamo dire che:

(4) per ogni x esiste z con $x \leq z$, anzi, z è strettamente maggiore di x, e inoltre *non c'è niente in mezzo*
(5) $0 \leq x \quad$ per ogni x
(6) non esiste un massimo.

Benché abbiamo scritto $x = y$, il segno "=" non è lo stesso che $\leq$: è una proposizione diversa; anche se non lo esplicitiamo tra i segni della nostra struttura, qualche volta si sottointende. Comunque certamente la presenza di "=" ci permette di dire più cose; avere "=" tra i segni a disposizione modifica le cose che si riescono a dire. In generale c'è "=" nel 99% dei casi. Però non è un assoluto, non è una necessità, è una scelta di convenienza, e quindi qualche volta se non è conveniente non si mette. Immaginiamo di avere un $\leq$, e scrivere tutto solo con $\leq$, definendo "=" in questo modo:

$$x = y \equiv_{def} x \leq y \;\&\; y \leq x$$

Comunque sia, noi supponiamo di avere anche =.

Si noti che dire che vale 1. vuol dire che vale $\forall x(x \leq x)$. E x dove varia? In questo caso in $\mathbb{N}$, ma potrebbe essere un altro dominio, un D generico:

1*. $\forall x \in D\,(x \leq x)$

2*. $(\forall x, y, z \in D)(x \leq y \;\&\; y \leq z \to x \leq z)$ (dove abbiamo usato la scrittura $\forall x, y, z \in D$ per dire $\forall x \in D, \forall y \in D, \forall z \in D$): questa è la scrittura in una sola proposizione di 2.

Analogamente,

3*. $(\forall x, y \in D)(x \leq y \;\&\; y \leq x \to x = y)$

è una riscrittura di 3. Ora passiamo a 4.: come esprimere meglio quel "*non c'è niente in mezzo*"?

Dati x e z, *non c'è niente in mezzo* vuol dire che qualunque y si supponga sia in mezzo, o è x o è z. Scritto in simboli:

4*. $\forall x \exists z(x \leq z \;\&\; z \neq x \;\&\; \forall y(x \leq y \;\&\; y \leq z \to y = x \vee y = z))$

dove $\neq$ vuol dire semplicemente $\neg(z = x)$, è un'abbreviazione per la negazione di uguaglianza.

È conveniente definire il "minore stretto" ponendo: $x < y \;\equiv\; x \leq y \;\&\; x \neq y$ Allora 4*. equivale a:

$$\forall x \exists z(x < z \& \;\forall y(x < y \& z \leq y \to y = z))$$

Abbiamo poi (5) che è chiaro:

5*. $\forall x(0 \leq x)$

Si noti che qui stiamo assumendo che nel nostro alfabeto ci sia un elemento distinto 0 su cui poniamo tale condizione. Se però non abbiamo un segno 0 potremmo ugualmente dire la stessa cosa, ma in questo modo:

5**. $\exists w \forall x(w \leq x)$

e con ciò diremmo che c'è un primo elemento, più piccolo di tutti. Naturalmente questo primo elemento risulta automaticamente unico perché se anche $w' \leq x$ per ogni x, segue che $w' \leq w$ e $w \leq w'$, e quindi sono uguali.

Il compito che ci siamo prefissati è, come per le parentele, cercare di descrivere al nostro meglio l'ordine di questa specifica struttura e ci stiamo riuscendo. Dopo aver detto che c'è uno 0, diciamo che *non esiste un massimo* (6)... anzi questo lo abbiamo già detto. Perché già detto? Come potremo scrivere esplicitamente *non esiste un massimo*? Con $\forall x \exists z(x < z)$, che dice appunto che non c'è minimo, in quanto per ogni x c'è uno z strettamente maggiore. L'abbiamo già detto perché è dentro la condizione 4*, nel senso che da 4* segue $\forall x \exists z\,(x < z)$. La proposizione 4* è del tipo: $\forall x \exists z$ applicato ad una proposizione di tipo $Axz\&Bxz$, e da questo segue che $\forall x\exists zA(x,z)$, perché il sequente

$$\forall x\exists z(A(x,z)\&B(x,z)) \vdash \forall x\exists zA(x,z)$$

è derivabile. Lo vediamo, come esercizio di deduzione. Chiaramente $A\&B \vdash A$ è derivabile. E poi:

$$\cfrac{\cfrac{\cfrac{\cfrac{A(x,z)\&B(x,z) \vdash A(x,z)}{A(x,z)\&B(x,z) \vdash \exists zA(x,z)}}{\exists z(A(x,z)\&B(x,z)) \vdash \exists zA(x,z)}}{\forall x\exists z(A(x,z)\&B(x,z)) \vdash \exists zA(x,z)}}{\forall x\exists z(A(x,z)\&B(x,z)) \vdash \forall x\exists zA(x,z)}$$

Si noti l'ordine delle regole, in modo da poter osservare le condizioni sulle variabili.

Ma si poteva fare anche così: non esiste uno z tale che per ogni x, $x \leq z$, cioè $\neg\exists$. Cambiamo lettera, mettiamo w, in modo che non ci confondiamo come intuizione ed avremo: non esiste w che sia maggiore di tutti gli x

6*. $\neg\exists w\forall x\,(x \leq w)$

questo(6*) segue da (4*)? Vogliamo dimostrare che (6*) segue da (4**) cioè:

$$\forall x\exists z(x < z) \ \vdash \ \neg\exists w\forall x(x \leq w)$$

Intuitivamente, questo è ovvio che valga: se, qualunque sia x c'è qualcosa maggiore di essa, non può esserci un w maggiore o uguale a tutti gli elementi del dominio, e basterà prendere qualcosa che è sopra quel w come dice (4**). L'idea è: se esistesse un w tale che $\forall x(x \leq w)$, cosa potremmo dire? Dalla pura esistenza $\exists w\forall x(x \leq w)$ vogliamo dedurre qualcosa e per far ciò assumiamo $\forall x\exists z(x < z)$ e vogliamo giungere ad una contraddizione. Allora scriviamo $\forall x\ (x \leq w)$. Innazitutto, prima di fare i conti, ragioniamo, perché il formalismo segue il ragionamento - noi riusciamo a chiarire quale è il ragionamento. Il ragionamento era: supponiamo che valga *per ogni x esiste uno z più grande*, se questo è vero allora non può esistere un w che sia più grande di tutti. Noi invece assumiamo che esista un elemento w più grande di tutti, e applichiamo il fatto che ne esiste uno strettamente più grande. Così facendo otterremo che non è vero che sono tutti $\leq$ di w. Adesso proviamo a scriverlo in formule.

Applichiamo $\forall x\exists z(x < z)$ a w, e otteniamo uno z tale che w è minore di z: $\exists z(w < z)$, dove w era quello che si assumeva esistere. Avremo: $\forall x\ (x \leq w)$. Dov'è la contraddizione? Proprio in quell'"*esiste uno z che è maggiore di w*", mentre ora ci ritroviamo a dire "*ogni x è* $\leq$ *w*". Come si fa ad ottenere questa contraddizione? Qualunque sia lo z che esiste per cui $w < z$, si applica $\forall x\ (x \leq w)$ a tale z, che ci dà $x \leq z$. Ed ecco la contraddizione da $w < z$ e $z \leq w$. Perché $w < z$ per definizione era $w \leq z\ \&\ w \neq z$; $z \leq w$ e $w \leq z$ ci danno $w = z$ che con $w \neq z$ ci dà una contraddizione:

$$\cfrac{\cfrac{w \leq z, z \leq w \vdash z = w \quad z = w, z \neq w \vdash \bot}{w \leq z, z \neq w, z \leq w \vdash \bot}}{w \leq z\&z \neq w, z \leq w \vdash \bot}$$

La derivazione allora continua con:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{w < z, z \leq w \vdash \bot}{w < z, \forall x(x \leq w) \vdash \bot}}{\exists z(w < z), \forall x(x \leq w) \vdash \bot}}{\forall x \exists z(w < z), \forall x(x \leq w) \vdash \bot}}{\forall x \exists z(w < z), \exists w \forall x(x \leq w) \vdash \bot}}{\forall x \exists z(w < z) \vdash \neg \exists w \forall x(x \leq w)}$$

Un'altra osservazione: queste tre proposizioni (1*), (2*), (3*) e le altre due (4*), (5*) valgono nella struttura che abbiamo scelto. Ma valgono anche in tante altre strutture, diverse. Se togliamo le due condizioni (4*), (5*), allora ci sono tantissime strutture (X, R) in cui X è un insieme e $R(x, y)$ una relazione binaria, ovvero una funzione poposizionale con due argomenti $x, y \in X$, che soddisfi (1*), (2*), (3*) (scritte con $R(x, y)$ al posto di $x \leq y$). Ogni tale struttura si dice un insieme parzialmente ordinato; di solito si usa $\leq$, ma scrivere $R(x, y)$ invece che $x \leq y$ non cambia la sostanza.

Ma anche se lasciamo le condizioni (4*), (5*), possiamo trovare una struttura che le soddisfa tutte, e che è molto diversa da $\mathbb{N}$.

Dato un qualunque insieme finito di segni $W = \{s_1, s_2, \ldots s_k\}$ definiamo l'insieme W^* delle parole, o liste su W. Diciamo che:

$$nil \in W^*$$

$$\frac{a \in W^*}{a * s_i \in W^*} \quad \text{per ogni } i = 1, 2, \ldots, k.$$

Si dice che W^* è definito induttivamente (vedi seguito); le regole dicono che la parola nulla *nil* è un elemento, e che poi, a partire da ogni parola a si può ottenere una nuova parola componendo (con l'operazione $*$) con un qualunque segno dell'alfabeto W. Definiamo un ordine parziale su W^* dicendo che $x \leq z$ se z prolunga x, ovvero x è un segmento iniziale di z, cioè se partendo da x componendo con un qualche numeo di segni si può ottenere z.

È immediato vedere che $\leq$ così definito è davvero un ordine parziale. Ma si vede che soddisfa anche (4*), (5*). Infatti, è facile vedere che per ogni x, e ogni s_i si ha $x < x * s_i$. Ed è immediato che $nil \leq x$, per ogni parola x.

Allora con ciò abbiamo fatto un'altra scoperta: credevamo di descrivere $\mathbb{N}$, e invece non ci siamo riusciti: abbiamo appena descritto un'altra struttura, che soddisfa le stesse proposizioni che abbiamo scelto.

C'era un modo per ovviare questo inconveniente imponendo che

$$6^{**}. \quad \forall xy(x \leq y \ \vee \ y \leq x)$$

cosa che nelle parole non vale, e quindi anche aggiungendo questa proprietà $\mathbb{N}$ continua a essere un'interpretazione delle condizioni (4*), (5*) e (6**), mentre W^* non soddisfa (6**)

Ed ecco che abbiamo trovato un modo per specificare meglio che intendevamo parlare dei numeri naturali. Possiamo specificare così bene da avere solo i numeri naturali? Solo i numeri naturali sarebbe impossibile, eventualmente sarebbe a meno di isomorfismo. E questo, nel caso specifico dei numeri naturali, è sì possibile, però richiede un liguaggio simbolico più potente di quello che stiamo costruendo. In genere, se si prendono delle condizioni, ci sono tante interpretazioni che soddisfano quelle condizioni, e vedremo che questo non è uno svantaggio. E comunque così stanno le cose.

La fatica che si fa in questi casi è di pensare per la prima volta, in modo consapevole, al legame che c'è fra una *condizione* e *il luogo in cui viene soddisfatta*, mentre di solito questo lo si considera come dato. Ora stiamo vedendo che è una cosa che si può studiare, laddove solitamente ci vien detto: questo è un gruppo, questo è un ordine, etc... senza specificare per bene cosa sono le condizioni e cosa le strutture in cui valgono.

2.2. Strutture algebriche.

Gruppi. Che cos'è un gruppo? La definizione usuale recita: un insieme con una operazione binaria associativa $+$ e un elemento neutro e, e in cui ogni elemento ha opposto.

Vediamo ora come si può esplicitare per bene. Dire che $+$ è un'operazione binaria su X significa che è una funzione con due argomenti in X e risultato in X, cioè $+ : X, X \longrightarrow X$. Spesso si chiede che "$X$ sia chiuso per l'operazione $+$", ma questo è incluso nella richiesta che $+$ sia una funzione da X, X verso X.

Dire che $+$ è associativa significa che vale:

$$\text{per ogni } x, y, z \in X, \quad x + (y + z) = (x + y) + z$$

e dire che e è neutro significa che vale:

$$\text{per ogni } x \in X, \quad x + e = x \;\text{ e } e + x = x$$

e dire che ogni x ha opposto significa che vale:

$$\text{per ogni } x \in X, \text{ esiste } y \in X \text{ con } x + y = eey + x = e$$

Ora analizziamo per bene il significato di questa definizione, dal punto di vista della logica. Abbiamo un linguaggio con due segni particolari: uno, il segno $+$, è un segno per una operazione binaria, e un segno e per un elemento costante. Una interpretazione di questo linguaggio, è data da un insieme, con una operazione binaria e con una costante, che vuol dire un elemento di quell'insieme. Questo significa interpretare questo linguaggio in una struttura specifica che è data da un'operazione e da un elemento neutro. Quando consideriamo un gruppo, quel che facciamo è prendere un insieme specifico, un'operazione specifica e un elemento dell'insieme specifico e dire quell'insieme con quell'operazione con quell'elemento soddisfa gli assiomi della teoria dei gruppi, che ora scriviamo più formalmente con:

$$\forall xyz(x + (y + z) = (x + y)z)$$
$$\forall x(x + e = x \& e + x = z)$$
$$\forall x \exists y(x + y = e \& y + x = e)$$

Tutto ciò significa che tale struttura (l'insieme con operazione con elemento) è *un modello della teoria dei gruppi*.

Come sarà la teoria dei gruppi commutativi altrimenti detti abeliani? Bisognerà aggiungere: $\forall x, y(x + y = y + x)$.

Anelli. . . . e per la teoria degli anelli? Bisogna aggiungere un'operazione ed allora otterremo un linguaggio con le operazioni $+$, $-$ e $\cdot$, una costante 0 e un'altra costante 1.

Per gli anelli, il linguaggio ha i segni: $+$ op. binaria, $\cdot$ nuova op. binaria, $=0$ costante e1 nuova costante. Dobbiamo richiedere:
per l'operazione $+$
- associativa
- neutro: $x + 0 = x = 0 + x$
- opposto:
per l'operazione $\cdot$
- associativa
- neutro: $x \cdot 1 = x = 1 \cdot x$
- distributiva: $\forall xyz(x \cdot (y + z)) = x \cdot y + x \cdot z$
  e dall'altra parte $\forall xyz((y + z) \cdot x = y \cdot x + z \cdot x)$.

Si noti che, per essere pignoli, bisognerebbe scrivere non i segni $\cdot$ e $+$, ma i segni del linguaggio che possono essere diversi: la $+$ potrebbe essere una f e $\cdot$ una g. Perché insistere su questo? Per ricordare che questi sono solo dei segni le cui proprietà stiamo chiarendo in questo momento; usare un segno familiare come $+$ può indurre a credere di conoscere già le sue proprietà, mentre sappiamo che esse non sono automatiche. Si ricordi che stiamo istruendo il robot e il robot non sa nulla. Quindi dobbiamo dire tutte le proprietà di una cosa, anche quelle che ci sembrano ovvie!

Ebbene, questo è un anello. E un campo? Si deve aggiungere: per ogni x diverso da 0, esiste l'inverso:

$$\forall x(x \neq 0 \to \exists y(x \cdot y = 1 \,\&\, y \cdot x = 1))$$

Non c'è da aver paura a scrivere come al solito $x \neq 0$, che leggiamo "x diverso da 0", purché sia chiaro che questo abbrevia la proposizione $\neg(x = 0)$.

Un'osservazione tipica da logico: tutte le operazioni, le condizioni associativa-neutro, associativa-neutro si scrivono con equazioni, sono equazioni che valgono:
- $x \cdot (y \cdot z) = (x \cdot y)z$
- $x \cdot e = e$
-$e \cdot x = x$
sottintendendo "quantificazione universale". Possiammo fare lo stesso per l'opposto? Vediamo la prima parte dell'anello
- $\forall x \exists y(x + y = 0)$ questa è un'equazione? Che cosa intendiamo quando diciamo "per ogni x esiste un y tale che $x + y = 0$"? Una proposizione il cui senso è legare queste due variabili: per ogni x esiste un y dipendente da quell'x. Quindi questa non è un'equazione nel senso che dicevamo prima, non possiamo scriverla sottointendendo una quantificazione universale. Qui abbiamo bisogno di dire "per ogni x esiste un y". E per trasformarla in equazione? Semplicemente cambieremo lo stock dei segni a disposizione. Si aggiunge un'operazione unaria $-$ nella notazione con il $+$ e si scriverà:

$$x + (-x) = 0 \text{ e } -x + x = 0.$$

E quindi le operazioni sono:
  $x + (y + x) = (x + y) + z$
  $x + 0 = 0$

$0, x = x$
$x + (-x) = 0$
$-x + x = 0$
Si deve allora dire che un gruppo non è più una terna, ma una quaterna e allora si dovrà dire: $(X, +, -, 0)$ dove $-$ è un'operazione unaria.

Ecco allora la differenza: se noi addottiamo un linguaggio diverso, la classe dei gruppi (e degli anelli) diventa assiomatizzabile con pure equazioni; in algebra si chiama *classe equazionale*.

Questi brevi cenni solo per mostrare come la logica c'entra con l'algebra; in un certo senso l'algebra è un pezzetto della logica, quel pezzetto che studia le strutture fatte con insiemi e operazioni soltanto: qui infatti abbiamo la relazione $=$ e solo operazioni, solo funzioni. Ed, inoltre, abbiamo visto che scrivendo in modo diverso, si ottengono informazioni diverse, e cioè in questo caso che la classe è equazionale.

Campi. Vediamo le proprietà di campo. Un campo è un anello più o meno commutativo e ha inverso per ogni elemento non nullo. La classe dei campi è equazionale? Si possono trasfromare gli assiomi in equazioni? Dobbiamo prendere una funzione non definita per 0 (la chiamiamo $()^{-1}$). Come facciamo a esprimere il fatto che una operazione non definita su 0? Dobbiamo cambiare completamente il concetto di linguaggio in cui i segni non rappresentano più funzioni ma funzioni parziali e questa però è una complicazione. Se vogliamo ottenere le funzioni nel senso attuale che siano definite dappertutto, non possiamo adottare questo sistema. E allora si può dire che la classe dei campi è equazionale, cioè la si può raccontare con qualche equazione? Se la soluzione di considerare la funzione $()^{-1}$ non la vogliamo adottare perché non è definita dappertutto, quale può essere un'altra via d'uscita?
I logici dicono semplicemente che non c'è via d'uscita. La classe dei campi non è equazionale e non c'è modo di raccontare l'esistenza dell'inverso su elementi non nulli se non con una formula che usa un "per ogni" e un "esiste".

Tutto ciò non fa parte della logica pura, però illustra molto bene che considerazioni intorno al linguaggio si possono fare, quali sono le sue interpretazioni, che cosa si può dedurre, etc.; e che i legami di queste cose con strutture matematiche come sottogruppi, sottoalgebre sono rilevanti. Più avanti cercheremo di approfondire il perché.

2.3. Geometria analitica. Consideriamo ora la geometria analitica. Che cosa può dirci sulla geometria analitica quel che abbiamo fatto finora? Intanto cerchiamo di capire che tipo di linguaggio c'è. Possiamo semplicemente pensarla come $\mathbb{K} \times \mathbb{K}$ dove $\mathbb{K}$ è un campo e, se proprio vogliamo, possiamo scrivere pure $\mathbb{R} \times \mathbb{R}$, intendendo considerare il piano come appare in figura 1.

Scriveremo ora solo coppie di numeri, immaginandole rappresentate su quel piano. Dobbiamo solo dire cosa sono i numeri reali $\mathbb{R}$. Le definizioni usuali di numeri reali o sono assiomatiche, come quelle che stiamo considerando qui noi, o sono costruzioni sul tipo di quelle introdotte da Cantor o Dedekind, che utilizzano un concetto di insieme infinito come qualcosa di *dato*, come oggetto che diventa elemento di qualcos'altro, considerandolo,

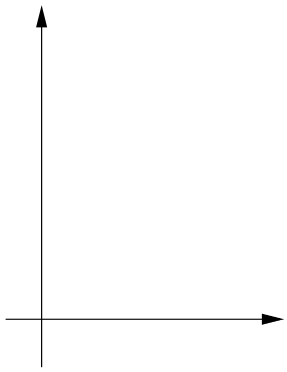

FIGURA 1. piano reale

quindi, come qualcosa di essenziale, ineliminabile. Basti pensare alle sezioni di Dedekind: sono una partizione dei numeri razionali che è esplicitamente infinita, che diventa un elemento dei numeri reali. Sicché, dal punto di vista dell'analisi costruttiva dei fondamenti, quelle definizioni non sono affatto una cosa semplice, anche se le usiamo in continuazione.

Parlare di *geometria analitica* vuol dire parlare di $\mathbb{R} \times \mathbb{R}$. Spiegando in che modo, arriviamo al concetto di insieme definibile. Supponiamo di avere un certo linguaggio, un certo alfabeto, chiamiamolo $\mathcal{L}$ - ma lo definiremo meglio - e supponiamo che $A(x)$ sia una proposizione in quel linguaggio, una formula - la definiremo bene. Supponiamo di interpretare tutto ciò in una struttura, costituita da un un dominio D e per ogni segno del linguaggio un qualcosa che lo rappresenti in D. I segni del linguaggio sono sostanzialmente funzioni e relazioni, perché i connettivi sono trattati come abbiamo visto finora. Vediamo cosa accade nel caso specifico dei numeri reali. Supponiamo di avere una proposizione scritta nel linguaggio dei numeri reali e che parla di x, y, ad esempio $x+y=0$. Questa è una formula (così la si chiama di solito, e anche noi lo chiameremo in questo modo quando avremo la definizione induttiva del concetto di formula), e $x+y=0$ è una funzione proposizionale su x, y che può essere vera o no a seconda di quali valori diamo a x e a y. L'insieme delle coppie che rendono vera la funzione proposizionale è $\{(x, y) : x + y = 0\}$, e cioè l'insieme di tutte le coppie costituite da due numeri che sostituiti rispettivamente a x e a y danno come somma 0, è un sottoinsieme di $\mathbb{R} \times \mathbb{R}$ che si dice definito da questa funzione proposizionale. Si noti che possiamo considerare qualsiasi funzione proposizionale che abbia come variabili x e y. Ed allora in generale possiamo dire che *un sottoinsieme è definibile se c'è una funzione proposizionale che definisce quell'insieme.*
Nel caso di questo specifico esempio $x + y = 0$ definisce sottoinsieme la bisettrice del secondo e quarto quadrante (come in figura 2).
In altri termini, 'il sottoinsieme definito da' corrisponde esattamente al concetto di 'luogo geometrico dei punti' - usando, con ciò un linguaggio ormai desueto. Ma questa è logica. Se restiamo alle formule che sono espresse in forma di equazione, allora potremo ridurle sempre alla forma generale $p(x, y) = 0$ perché se mettiamo il segno d'equazione tra due polinomi

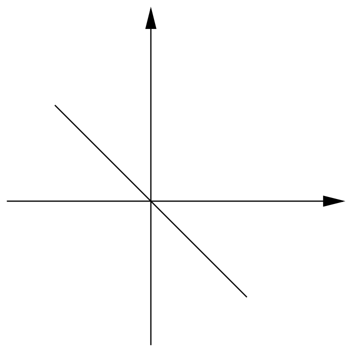

FIGURA 2. x+y=0

$p(x,y) = q(x,y)$, sappiamo che questo è come dire $p(x,y) - q(x,y) = 0$, che, allora, vuol dire che potremo ridurre tutte le formule atomiche con l'uguale tra due polinomi a casi particolari di $p(x,y) = 0$

E i connettivi, e i quantificatori che fine fanno? Lo sappiamo già perché quando diciamo "il luogo geometrico dei punti che soddisfano il sistema"

$$\begin{cases} p(x,y) = 0 \\ p'(x,y) = 0 \end{cases}$$

cosa facciamo se non *la congiunzione* di due funzioni proposizionali? Cosa se non prendere l'intersezione dei due insiemi definiti? E così arriviamo a quel che si dice *far sistema*, e cioè trovati quali sono gli (x,y) per cui vale $p(x,y) = 0$, e $p'(x,y) = 0$ e cioè trovato ciò che disegna $p(x,y) = 0$, si disegna poi $p'(x,y) = 0$ e, quindi, si fa l'intersezione.
E i quantificatori? Supponiamo di avere l'insieme definito dalla funzione proposizionale $A(x,y)$. Se scriviamo $\exists y A(x,y)$ in tal caso chiederemo di avere l'insieme $\{x : \exists y\ A(x,y)\}$ in cui compare soltanto l'insieme degli x in quanto y è stato quantificato; quella y, infatti, non è più una variabile libera perché è stata bloccata dall'azione del quantificatore esistenziale $\exists y$, continuando, invece, x ad essere libera. Questo vuol solo dire che ci interessano le x, i numeri reali che soddisfano $\exists y A(x,y)$, e quindi che non dobbiamo più operare sostituzione sugli y, sicché gli elementi che soddisfano $\exists y A(x,y)$ sono numeri reali, e non più coppie. E quali sono nel disegno? Mettiamo x nell'insieme quando c'è un y per cui la coppia (x,y) soddisfa A. Perciò l'insieme $\{x : \exists y\ A(x,y)\}$ è la proiezione sull'asse delle x (figura 3).

E il $\forall$? Se qui scrivessimo $\{x : \forall y\ A(x,y)\}$ cosa risulterebbe? Da A verrebbe il vuoto, nessun x, perché risulterebbe $A(x,y)$ per tutti gli y che vuol dire x lo mettiamo se tutta la retta (verticale) che passa per esso sta in A. Quindi, con questo A qui, nessuno!

2.4. Metodo assiomatico. E la geometria sintetica, quella di Euclide? Cosa c'entra con la logica? Qui daremo solo un accenno alla questione.

Si può dire che, in un certo senso, la geometria nasce insieme alla logica perché i greci scoprirono che le idealizzazioni della geometria (punti, rette, triangoli,...) si potevano trattare in modo rigoroso, e cioè usando solo logica e

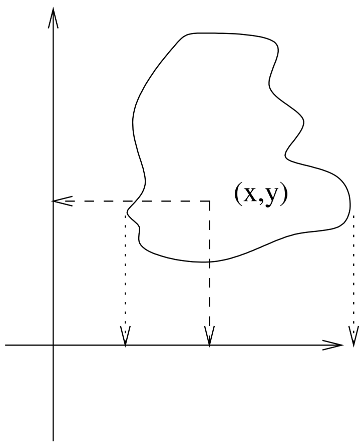

FIGURA 3. proiezione di A sull'asse x

alcuni postulati. Questo modo è il così detto *metodo assiomatico di Euclide*. Vediamolo.

Euclide visse circa nel 300 a.c. Nei suoi **Elementi** troviamo, raccolto, e scritto in un linguaggio geometrico, tutto quello che i Greci avevano imparato di matematica. Solitamente si crede che il libro tratti solo di geometria e invece c'è anche molta matematica (semplicemente quei capitoli che riguardano l'aritmetica li si saltano sempre perché oggi la sappiamo fare molto meglio, nel senso che abbiamo sviluppato una notazione algebrica molto più potente di quella che loro padroneggiavano). Gli **Elementi** di Euclide sono, dunque, tutta la matematica e non soltanto la geometria e sono organizzati in questo modo:
Le *definizioni*, ad esempio:

(1) punto è ciò che non ha parte
(2) linea è ciò che non ha larghezza (una lunghezza senza larghezza)
(3) etc. per retta e segmento

Sono definizioni che di solito lasciano perplessi gli studiosi o i filosofi della matematica perché non sono propriamente matematiche.
Come abbiamo imparato il concetto di punto? In seconda o terza elementare ci hanno insegnato la definizione data da Euclide, cercando di far scattare in noi il processo di idealizzazione che conduce da punti disegnati a punti nel senso geometrico, che sono ciò che non ha parte. Ma questo punto come ciò che non ha parte non esiste di per sé. Ci sono punti in natura? Tutti gli esempi che diamo di punto altro non sono che quel che noi abbiamo in mente, ovvero l'idealizzazione. Questa idealizzazione richiede fatica mentale, è un processo che ci porta ad un concetto di punto *per noi*, per la nostra mente. Lo stesso vale per il concetto di retta: per poter esemplificare e rappresentarci questo concetto si prendono spaghi che vengono tesi, senza tuttavia tener presente l'azione della forza di gravità per cui, di fatto, quel filo di spago non sarà mai perfettamente dritto.
I *postulati*

1. per due punti passa una e una sola retta (in realtà i due punti devono essere distinti altrimenti sappiamo che ce ne possono essere tante);
2. angoli retti sono uguali da qualunque parte si considerino;
3. l'angolo retto è metà dell'angolo piatto cioè di una retta;
4. dato un punto ed un segmento posso tracciare la circonferenza con quello come centro e il segmento come raggio;
5. data una retta e un punto fuori da essa esiste ed è unica la sua parallela per quel punto (questo, com'è noto è il più famoso)

Troviamo poi le *nozioni comuni* che sono, per esempio, la proprietà transitiva dell'uguaglianza (che egli esprime a parole) chiamandola uguaglianza di grandezze, con le quali si intendono sia segmenti, che aree, che triangoli, etc;, e altre proprietà.
Ed infine le *proposizioni* che sono gli enunciati della matematica.

E oggi? Il lavoro della matematica si è, nel tempo, sviluppato in questo modo: date le nozioni comuni, ottenere proposizioni a partire dai postulati e con l'aiuto delle definizioni che forniscono i nomi con cui usare i concetti che abbiamo prodotto e capito. Che significa ottenere? Sostanzialmente *partire dai postulati e usare le nozioni comuni.* E a cosa corrispondono oggi le nozioni comuni?
Precisamente alla *logica* o, per essere precisi, agli *assiomi logici* più *le deduzioni che si usano per ottenere le proposizioni*: quelle deduzioni sono, infatti deduzioni logiche. Si potrebbe benissimo tentar di scrivere tutta la geometria di Euclide usando solo postulati e logica, con quello che abbiamo visto fino ad adesso e niente di più.
Le proposizioni oggi sono i *teoremi*, nel senso che sono proposizioni dimostrate a partire da assiomi, i postulati di Euclide. Qualche volta si distinguono assiomi non logici da assiomi logici - questa è la terminologia; ma forse sarebbe più intuitivo continuare ad usare il termine "postulato", perché ricorda che cosa si fa, e cioè che si postula, si chiede che valga una certa cosa, quindi è specifica del tema, mentre assioma non sembra restituire l'idea di qualcosa che si considera sempre vero.
E le definizioni? Il metodo assiomatico moderno, non essendo capace di risolvere questo problema, ha spiegato che non costituiva problema: infatti, non esiste nemmeno una trattazione del concetto di definizione. E quindi? Vediamo il seguente schema

Euclide	Oggi
definizioni	/
postulati	assiomi (non logici)
nozioni comuni	logica (assiomi logici)
proposizioni	teoremi (proposizioni dimostrate)

Hilbert che fu il massimo rappresentante di questo approccio, scrisse nel 1899 i ***Grundlagen der geometrie***, testo fondamentale nella storia della matematica perchè è la prima trattazione completa della geometria secondo il metodo moderno che stiamo illustrando. Un episodio famosissimo racconta di quando Hilbert si trovò a discutere di queste cose con colleghi e allievi in birreria: "a me non importa se voi leggete, se intendete rette, piani, punti come boccali di birra, tavoli, o quel che volete. L'importante è che soddisfino quello che dico io". Con ciò Hilbert volle dire che qualunque interpretazione si dia del linguaggio, delle parole che compaiono, delle relazioni, etc., ebbene, devono soddisfare quello che si chiede all'interno del sistema, ovvero gli assiomi e nient'altro. Ora si tratterà di trovare un'interpretazione degli assiomi. Ma allora il concetto di retta e di punto saranno specifici di quella interpretazione, quindi non dando tanto una definizione generale, quanto semplicemente un concetto di punto, retta, etc. definito implicitamente dagli assiomi, nel senso che è qualunque cosa che soddisfi gli assiomi.

Le *definizioni* non si trovano nel metodo assiomatico moderno, perché si intende che gli enti di cui si parla siano definiti implicitamente dalle proprietà. Si prende un linguaggio del primo ordine, cioè un linguaggio in cui esprimere per esempio x *giace su* y:

$G(x, y)$ che vuol dire x *giace su* y

e poi si scrivono gli assiomi, qualcosa del tipo *per ogni due punti esiste un'unica retta*

$$(\forall x, y)(x \neq y \rightarrow (\exists! z)(G(x, z) \& G(y, z)))$$

E qui si procede con tutte le varie assunzioni sul predicato *giacere*, non dicendo mai che cosa sono i punti o le rette. Si usano le parole solite per aiutare l'intuizione, ma in realtà non si specifica che cosa significano. Quindi, qualunque interpretazione in cui vengano soddisfatti gli assiomi, poi va bene. E questo è il modo di procedere di oggi. Qual è il vantaggio? Beh, i vantaggi sono tanti, forse, anche gli svantaggi.

Un vantaggio è la maggiore generalità: non avendo un'interpretazione specifica in mente, si può parlare di tante cose, tante possibili interpretazioni nello stesso momento; e le proprietà derivano, solo tramite logica pura, dagli assiomi (logica pura nel senso di "solo formalmente", senza andare a vedere il significato, e cioè solo con deduzioni sulla forma delle proposizioni, indipendentemente dalla loro interpretazione).

Il vantaggio più importante è che una trattazione assiomatica è possibile anche quando l'intuizione vacilla: anche quando non abbiamo un'interpretazione intuitiva convincente, ciononostante il metodo assiomatico procede lo stesso. E facendo a meno dell'interpretazione, come abbiamo appena detto, e trattando le cose solo in modo formale. Ci bastano due soli esempi

(1) le geometrie non euclidee. Non è molto intuitivo che cosa vuol dire una geometria in cui ci sono tante rette parallele o nemmeno una, o in cui lo spazio è curvo. Ma dal punto di vista logico, metodologico è chiarissimo: è semplicemente il metodo assiomatico applicato a una geometria in cui vale una negazione del V postulato di Euclide. E geometria non euclidea risulta qualunque cosa soddisfi quella teoria. Ecco il salto: si parte da una geometria, che è quella piana di Euclide, e si cerca di parlarne, di descriverla. Constatando la forte problematicità di un discorso intorno ad essa, di una teoria per essa, si decide di tenerne ferma solo la descrizione, dimenticando tuttavia

l'oggetto descritto. E *la descrizione* diventa *la teoria*, anche se manca l'oggetto di cui questa teoria era la descrizione.

(2) L'apoteosi viene raggiunta con la teoria assiomatica degli insiemi: lì, infatti, abbiamo un'idea di insieme molto vaga, che fa scaturire in ciascuno di noi intuizioni diverse non facilmente comunicabili, e quello su cui ci si basa alla fin fine è solo la teoria assiomatica e cioè un concetto di insieme che soddisfi le proprietà in essa stabilite. La matematica oggi si basa su una teoria di questo tipo.

Quando parliamo di teoria intendiamo semplicemente un campo studiato facendo uso del metodo assiomatico moderno e cioè sostanzialmente scegliendo uno stock gli assiomi. Ecco, prima ancora di decidere che linguaggio scegliere (che primitivi usare nel linguaggio, ovvero che relazioni, che funzioni, che costanti) si deve decidere quali assiomi (che proprietà assumere su questi), e che logica scegliere (infatti possiamo scegliere tra logica intuizionistica, logica classica o logiche più deboli ancora). Una volta specificate queste tre cose, abbiamo quella che si chiama una teoria, o teoria assiomatica. E il metodo assiomatico consiste nel raccontare le proprie conoscenze dentro teorie assiomatiche e semmai ci fosse discussione, beh, si discutono gli assiomi perché nel momento in cui si specificano linguaggio e assiomi la discussione non ha più ragion d'essere perché, da quel punto in poi, è solo questione di deduzione logica.

APPENDICE

3. Sul principio della doppia negazione

Tornando l'altra volta, in aereo, non avevo niente di meglio da fare che leggere la rivista della compagnia aerea. C'era una lettera al direttore da parte del presidente dell'Associazione nazionale dentisti italiani (spero che tra voi non ci sia il figlio di un dentista). Evidentemente, in un precedente articolo si era parlato dei rischi connessi all'uso di mercurio nelle otturazioni, e allora questo presidente, rispondeva con le seguenti parole:

> "In nome e per conto dell'Andi, Associazione nazionale dentisti italiani, evidenzio che l'amalgama usata per l'otturazione delle lesioni della carie è considerata sicura ed efficace, non essendo, allo stato della scienza medica, dimostrato che la piccolissima quantità di mercurio utilizzata sia irreversibilmente dannosa per la salute".

Con una risposta simile, ve lo fareste voi mettere in bocca il mercurio o no? Certo che no, mi direte.

Ma allora non state facendo altro che osservare la differenza tra le due regole sulla negazione viste alla fine della lezione **??**. Il ragionamento del dentista è circa questo: non c'è prova che dica che faccia male, allora fa bene. Il fatto che faccia male è escluso, quindi fa bene. Quindi da un'informazione negativa si ottiene un'informazione positiva. Si usa la riduzione all'assurdo.

Mentre l'introduzione della negazione dice che se ho un'informazione negativa, posso solo ottenere ancora un'informazione negativa, e cioè: se da una certa assunzione si trova una contraddizione, allora quella certa assunzione è negata.

Se ci concediamo una minima digressione di commenti anche dettati da opinioni personali, questo ci permette di chiarire la differenza fra logica intuzionistica e logica classica. Ripartiamo dall'esempio del dentista. È come se dicesse: la scienza non ci ha ancora fatto vedere una contraddizione, quindi quella certa procedura si può usare. Questo vuol dire usare, avere un certo concetto di verità rigido, fisso, assoluto, ma anche completo, cioè che discrimina tutte le cose in buone e cattive. Questo ricompare nel principio del terzo escluso. Nel momento in cui si dice che se non si ha evidenza del fatto che valga la negazione, allora vuol dire che valeva la proposizione, si fa un riferimento preciso a questa discriminazione delle proposizioni in due: vere e false. E si assume anche, come il dentista, di avere accesso a questa verità o falsità. Il dentista precisamente dice: "io non ho ancora visto che faccia male, quindi fa bene"; dice quindi sia che vale $A \vee \neg A$ sia, e questo è più grave, che se valesse $\neg A$ lui lo saprebbe, e pertanto vale A.

Se ci pensate –d'ora in poi esprimo opinioni in modo più netto– questo è collegato ad un modo di intendere il mondo e la vita che poi porta a scontri irrimediabili, come fra Islam e libero mercato. Perché ciascuno dei due crede di avere la verità, e non lascia spazio alla verità dell'altro. Se invece quando si parla di verità ci si autoimpone, come si fa in logica intuzionistica, di presentare anche una argomentazione per quella verità, allora ci sarebbe più spazio per le opinioni, e di scontri frontali ce ne sarebbero di meno. Detto in altri termini, lo scontro a cui assistiamo negli ultimi mesi, sempre più irrisolvibile, a ben pensarci è dovuto allo scontro tra due modi di intendere la verità, entrambi considerati come assoluti. Da una parte libero mercato, da una parte una visione religiosa rigida; quel che non va bene è che ciascuno dei due ragiona su una verità prefissata e totalizzante. Siccome che non è vero che no, allora deve valere; tutto quel che non è in contrasto con la propria verità, automaticamente è anch'esso vero.

Credo che per ragionare così, in questo modo classico, si debba veramente avere una concezione rigida, molto lontana da noi, indipendente da quello che noi riusciamo ad ottenere, e quindi è questo che porta a risultati al di là di quello che noi vorremmo. Sembra che ci sia un legame tra l'assumere una logica più debole e una visione in cui, come dicevo, quando si dice qualcosa si deve anche portare un'argomentazione. Questo equivale a dire che se che c'è spazio per un dibattito, nel senso che se non si hanno argomentazioni né per A né per $\neg A$, su A resta spazio per la discussione. Faccio fatica a parlare di queste cose perché, siccome ho scritto qualche articolo su questa cosa, ho il timore di farmi trasportare.

Questo atteggiamento nei confronti della verità, che ho definito classico o rigido o assoluto, sembra indipendente da visioni religiose possibili. A quanto si vede, non solo oggi si ragiona pensando che ci sia una verità data, quella che è accessibile dalla scienza, ma anche che sia divisa in due. Da un lato ci sono le cose vere, da un lato le cose false, non c'è una via di mezzo.

Ci sono molti scienziati che si dichiarano scienziati proprio in quanto senza verità rivelata, che poi in realtà adottano questo atteggiamento. Per fare capire cosa intendo, posso dire che nella pratica mi è successo una volta di ascoltare una conferenza di uno studioso di meccanica quantistica, abbastanza noto in Italia, il quale faceva una discussione molto raffinata –che io non saprei ripetere– sulla logica di quel che è osservabile e di quel che non è osservabile. E alla fine diceva che per raccontare tutto aveva bisogno di un logica al livello del metalinguaggio che parli della verità. Diceva: "ne prendiamo una a caso, perché la verità è quello che è, e allora tanto vale prendere la logica classica". A quel punto io ho provato a osservare che nel momento in cui si assume la logica classica come metalinguaggio, questo comporta una serie di conseguenze anche nella proprie teorie specifiche. Ma non c'era modo di farmi capire; si direbbe proprio che non solo alcuni scienziati hanno fiducia in una concezione come quella descritta, ma addirittura non riescono a vederne un'altra. Ci sono vissuti talmente dentro che non riescono neppure a concepire un modo diverso di intendere la verità.

Vediamo su due esempi i due modi principali di intendere la verità. Sono due esempi molto simili. Si può dire, in modo positivo, che

1. Esistono gli extraterrestri.

Quindi questa proposizione sarebbe del tipo $\exists x A(x)$. Oppure si può dire:

2. Non è escluso che esistano gli extraterrestri $\neg\neg\exists x A(x)$,

Lo stesso schema c'è anche in:

3. Io Giovanni Sambin ho un antenato francese, che sarebbe del tipo $\exists x A(x)$ oppure

4. Non è escluso che io, Giovanni Sambin, abbia un antenato francese $\neg\neg\exists x A(x)$.

I due esempi sono molto simili. Allora vediamo il secondo, per mancanza di fantasia. Supponiamo che io dica: "Io ho un antenato francese, tanto che il mio cognome non si deve pronunciare Sambìn come dicono in Veneto, ma Sambèn perché è di origine francese, come Carden, come..." (per inciso Carden è Cardin, viene dalle mie parti anche lui). Allora, se dico: "Ho un antenato francese" cosa vuol dire? Cosa capite che io voglia comunicare con questa affermazione? La capite nel senso che io so chi è costui, da una documentazione storica che dice che risalendo il mio albero genealogico si arriva a questo tizio che era un francese. Per inciso se fate su Google la ricerca sulla parola Sambin, vedete che c'è un architetto francese con questo nome vissuto nel 1500 a Digione, Tolosa, non mi ricordo bene.

In realtà la situazione è che io non ho nessuna evidenza del fatto di avere un antenato francese, e tanto meno che sia quell'architetto, e quindi la situazione è la seconda. Questo non dice molto perché non esclude che sia un mio antenato, in quanto ha il mio cognome, non è escluso; non c'è modo per tracciare tutta la mia discendenza fino al 1500 e dire che quel signore non è mio antenato. Però, anche se non è escluso, non segue assolutamente che io so di avere un antenato francese. Vedete la differenza? Nella pratica, lasciate perdere che siete in un corso di logica, immaginate che siamo in un salotto, oppure al bar e ci siamo incontrati per la prima volta e vi dico: io ho un antenato francese e capite una cosa. Se vi dico: non è escluso che io abbia un antenato francese, capite un altra cosa. È una distinzione che si farebbe naturalmente.

Una cosa simile vale per gli extraterrestri. Anzi qui quantificazione possiamo pensarla sui corpi celesti, sui pianeti, su qualche galassia... e la proposizione dovrebbe essere: "In un qualche pianeta, di qualche sistema solare, di qualche galassia, si è sviluppata una vita che noi (naturalmente) consideriamo intelligente, tanto da poter dire che anche lì ci sono animali simili in qualche modo a noi".

Si vede allora che in questo esempio degli extraterrestri c'è un altro fatto interessante, e cioè che non è decidibile la proprietà: se in un qualche pianeta troviamo degli esseri viventi, dobbiamo considerarli intelligenti o no? Non sappiamo deciderlo nemmeno per gli scimpanzé, che abbiamo qui vicino, figuriamoci se dovessimo deciderlo per ipotetici abitanti di un ipotetico pianeta. Noi siamo ben sicuri solo in un caso, abbiamo solo una informazione positiva: se capita che un marziano con la sua astronave atterra a piazza San Pietro e dice: "siete tutti prigionieri", allora decidiamo sicuramente che quello era intelligente. Qui intelligente vuol dire che almeno ha una tecnologia superiore alla nostra. Ma in realtà di fronte a un pianeta specifico, con una specie di bestie specifica, non siamo nemmeno tutti d'accordo se dovremmo considerarli o no extraterrestri, nel senso che intendiamo quando si dice questo. Se su Marte si trova un po' di muffa o di batteri, diciamo che sono extraterrestri? No. Ma se arriva quello dell'astronave a Piazza San Pietro, sì. Ma non sappiamo decidere esattamente dove sta il mezzo. Detto in altri termini, non abbiamo un modo per scandire tutti i pianeti e per ciascun pianeta decidere se lì c'è vita intelligente o no.

Talvolta si assume che ci sia un numero infinito di galassie, quindi un numero infinito di pianeti. Ma se anche fosse che l'universo è finito, praticamente dobbiamo considerarlo come infinito, perché non sappiamo decidere per ciascun pianeta se ci sono o no degli esseri viventi da chiamarsi extraterrestri. Quindi comunque resta una grossa differenza tra l'affermazione positiva "esistono gli extraterrestri" e quella negativa "non è escluso che esistano"; la differenza è proprio quella in cui stiamo noi ora, dato che non sono ancora arrivati in Piazza San Pietro, ma non abbiamo modo di escluderlo.

Pensiamo ad i miei antenati che mettono meno ansia. Allora, i miei antenati sono certamente un numero finito, perché, per chi crede, male che vada si arriva fino ad Adamo ed Eva, per chi non crede, male che vada si arriva agli scimpanzé, 5 o 10 milioni di anni fa. Certamente 10 milioni di anni fa io non avevo antenati. Cioè i miei antenati si fermano prima, quindi essi sono in numero finito, enorme ma finito. Ma da questo non segue assolutamente che io posso scandire tutti gli esseri umani, e nemmeno in particolare i miei antenati, e vedere se ce n'è uno francese o no. Anche qui quindi rimane la differenza tra le due asserzioni.

Un conto è l'affermazione: ho un antenato francese. Possiamo esprimerla con qualcosa del tipo: esiste un individuo che è antenato di Giovanni Sambin, ovvero $\exists x(A(x, gs)\&Fr(x))$ dove $A(x, gs)$ sta per "x è antenato di G. S." e $Fr(x)$ sta per "x è francese". Un altro conto è dire $\neg\neg\exists x(A(x, gs)\&Fr(x))$, o equivalentemente (come vedremo negli esercizi) $\neg\forall\neg(A(x, gs)\&Fr(x))$. Anche se il dominio dei miei antenati è finito, anche se arrivo a dimostrare che non per tutti i miei antenati posso dire che non sono francesi, non mi è permesso concludere che ce ne è uno che lo è. Il fatto che non escludo che esista, è la situazione reale: io non ho alcuna evidenza che mi permetta di dire, ad esempio, che tutti i miei antenati vengono dall'Albania, e quindi non sono francesi. Analogamente, non abbiamo alcuna informazione positiva, alcuna evidenza che ci dica che in tutti i pianeti, tutti se anche finiti, non c'è un essere intelligente che sia uno. Non abbiamo informazioni di questo tipo. Ma non possiamo concludere che allora esiste un antenato francese, oppure che esistono gli extraterrestri. Perché anche se il dominio è finito, non abbiamo il modo di scandire tutti i casi, di vedere tutti i pianeti, e inoltre non abbiamo modo per ciascuno di trovare se ci sono o non ci sono vite intelligenti.

Vediamo la stessa cosa nell'altro verso. Se avessimo una proprietà $P(x)$, con x che varia su un dominio finito D, costituito ad esempio dagli elementi $d_1, \ldots, d_{451}$, e se inoltre fosse che sappiamo per ciascun elemento d di D decidere se vale o non vale $P(d)$, allora dalla affermazione $\neg\neg\exists xP(x)$ potremmo sensatamente concludere anche che $\exists xP(x)$. Il metodo per derivare la seconda dalla prima sarebbe questo: si fa una lista di tutti gli elementi, si prova per ciascuno se vale o non vale P. Non può succedere, per via della prima affermazione, che per tutti i $d \in D$ troviamo che vale il caso $\neg P(d)$. Allora in qualche parte abbiamo trovato che deve essere $P(d)$, e da questa informazione possiamo concludere che $\exists xP(x)$. Ma è da notare che questo metodo non può funzionare se cade anche solo una delle condizioni: 1. che il dominio sia finito, e 2. che sia decidibile per ciascun elemento d se vale P o no sudi lui. Formalmente, dire che P è decidibile su D vuol dire che vale $(\forall x \in D)(P(x) \vee \neg P(x))$.

Vi faccio un altro esempio, dello stesso tipo ma più semplice. È quello che abbiamo già visto, di quel dialogo: "vieni a cena domani? Non dico di no". Quello che chiede "vieni a cena" non ha modo di sapere se chi risponde dirà di sì o dirà di no... Beh, un modo ci sarebbe, ed è quello di aspettare fino alla sera dopo, e vedere se viene o no. Ma senza questo criterio di verifica, troppo prosaico, si resta nel dubbio; un conto è però sapere che la questione verrà decisa, un conto è avere una informazione positiva ora, sia in un verso sia nell'altro.

Nel caso della cena, la differenza tra le risposte “non dico di no” e “sì” è chiara, perché emotivamente ci immedesimiamo nella situazione; il fatto che domani verrà comunque deciso non modifica il dubbio di oggi. Ma anche per gli extraterrestri rimane la differenza tra “esistono gli extraterrestri” e “non è escluso che esistano”, perché qui non c’è nemmeno il modo di aspettare per vedere quale delle due vale. Per decidere, dobbiamo avere una informazione positiva, adesso. Se dico che esistono gli extraterrestri, vuol dire che adesso dico che esistono, e ne ho una prova. Se dico che non esistono, pure. Notate che qui con gli extraterrestri, allo stato attuale, siamo nella situazione seconda, in cui non sappiamo se esistono o no. Uno scienziato che dice che è escluso che esistano, fa lo stesso errore del presidente dei dentisti, perché dice: “con la scienza di oggi non li abbiamo ancora trovati, quindi non esistono”; questa non è però una prova, ma solo la mancanza di controprove.

Non abbiamo visto ancora che il mercurio fa male, quindi fa bene! (Per inciso, quando conversando ho raccontato questo, l’interlocutore mi ha ricordato che anche 30 anni fa dicevano che la pillola non faceva male, e invece adesso si comincia a vedere se non altro che le pillole di 30-40 anni fa forse non facevano male, ma che comunque non erano così innocue, quindi non era il caso di usarle allegramente).

Può essere che quello che si considera oggi come dimostrato in realtà con l’aumentare dell’informazione risulti falso. Se aumenta l’informazione scientifica sul mercurio, il dentista è costretto a ricredersi, a dire: “ok, mi ero sbagliato”. Mentre se si dimostrano le cose solo con argomentazioni positive, non si è costretti a ricredersi, perché è previsto lasciare spazio al dubbio.

Ora vediamo un po’ di storia, che può essere utile a capire il perché della situazione attuale. Chi ha cominciato a parlare di queste cose è un certo Brouwer. Come data di inizio si può prendere il 1907, la data della sua tesi, in cui Brouwer comincia a costruire una matematica, che si chiama matematica intuzionistica, che non usa mai il principio del terzo escluso o i ragionamenti per assurdo.

Purtroppo, la storia successiva non è molto divertente, anzi, è in qualche modo drammatica, perché quando sembra potersi sviluppare questa matematica costruttiva, si crea uno scontro accesissimo tra Brouwer e Hilbert, il famosissimo matematico tedesco, e nessuno dei due fa bella figura.

Ma per capire la storia, bisogna prima fare un passo indietro, e ricordare almeno con un cenno la storia della teoria degli insiemi e dell’infinito, che comincia con Cantor e che raccontano ormai da tutte le parti. Cantor comincia a studiare l’infinito in matematica e vede che si può trattare di infinito in termini matematici. Non entro nella discussione adesso, perché non ho tempo. La visione di Cantor era molto platonista, secondo lui la matematica esiste come esistono gli angeli o simili, in qualche mondo iperuranico a cui non abbiamo accesso, però ciononostante è una verità che non si discute, una verità assoluta.

Un vantaggio di questa teoria degli insiemi di Cantor è che permette di dare un linguaggio comune alla matematica, a tutta la matematica nota fino a quel momento e da quel momento in poi. Cantor comincia nel 1872-1873 a parlare di insiemi e del concetto di insieme. Pochi anni dopo Hilbert, alla fine dell’800, comincia a sfruttare la teoria degli insiemi di Cantor, compreso anche il suo aspetto platonistico, nel senso che Hilbert comincia a sfruttare l’idea che esistono in qualche parte degli insiemi infiniti, e questo gli permette di ottenere nuova matematica, in particolare nuove dimostrazioni. Ad esempio dimostrazioni di un enunciato del tipo: ogni spazio vettoriale ha una base, e anche altri teoremi più profondi in algebra e matematica in generale. Questi vengono dimostrati da Hilbert usando la teoria degli insiemi di Cantor, incluso il concetto di infinito attuale. Questo significa che anche gli insiemi infiniti vengono intesi come oggetti, completamente determinati, che esistono da qualche parte, di per sé, indipendentemente dal fatto che noi li pensiamo o meno. La matematica in cui viviamo anche oggi è quella sviluppata su questa base da quel momento in poi.

Brouwer è molto più giovane, e quando nel 1907 scrive la sua tesi ha 26 anni; quindi praticamente comincia a parlare appena può, e in termini storici viene subito dopo Cantor, perché per la storia 30 anni è come dire subito dopo.

Brouwer mette in discussione questa concezione platonista della matematica e in particolare degli insiemi infiniti. Si accorge del fatto che non possiamo ragionare così e nello stesso tempo mantenere l'interpretazione usuale della logica. Se si considerano anche insiemi infiniti, e si mantiene la logica classica, si perde la qualità dell'informazione in certe proposizioni matematiche. Ad esempio, per quella che nominavo poco fa: ogni spazio vettoriale ha una base, la dimostrazione data da Hilbert non ci permette in generale di costruire la base sul serio.

La proposta di Brouwer è di distinguere le proposizioni in cui abbiamo un modo per costruire quel che affermiamo che esista, da quelle in cui non ce l'abbiamo. Praticamente, in termini di logica, Brouwer dice che dobbiamo distinguere un'affermazione della forma "esiste" da una della forma "non è escluso che esista": mentre queste due forme possono in un certo senso considerarsi equivalenti quando ci restringiamo al caso finito e in cui inoltre supponiamo di poter decidere tutte le proposizioni, nel caso infinito certamente questa equivalenza non vale (come abbiamo visto per gli extraterrestri e gli antenati).

Notate che fino a quel momento, fino a Cantor non si era mai parlato esplicitamente di insiemi infiniti. Quando Aristotele parla di logica, uno dei principi più importanti è quello del terzo escluso (o altri equivalenti) perché di fatto continua a pensare a domini finiti e a proprietà decidibili. La diversità viene proprio quando si inizia a parlare di infinito; quindi Brouwer, in un certo senso, non è altro che la risposta storicamente immediata all'introduzione dell'infinito in matematica.

E, vista nello sviluppo storico, la sua osservazione è molto semplice e ragionevole: nel momento in cui si parla di infinito, non vale più il principio del terzo escluso o della doppia negazione, se si vuole mantenere la qualità dell'informazione, per cui un enunciato che afferma l'esistenza di qualche cosa deve anche permettere di produrre un testimone, un esempio di quello esiste. Detto in altri termini, se si vuole mantenere la distinzione fra due affermazioni come "ho un antenato francese" e "non è escluso che abbia un antenato francese", una distinzione che si fa anche in una conversazione da salotto, dobbiamo usare una logica diversa da quella classica, dobbiamo usare la logica intuzionistica.

Sia Brouwer sia Hilbert erano matematici di altissimo valore. Per darne un'idea, mi permetto di raccontare con parole mie un episodio importante della vita di Brouwer. Pur rimanendo fedele al suo progetto di sviluppare la sua visione alternativa della matematica, Brouwer ad un certo punto si rende conto che in quel modo non avrebbe mai fatto carriera. Allora molto semplicemente si propone: adesso dimostro un pò di teoremi, in modo da vincere una cattedra, e poi ricomincio a studiare alle cose che mi premono. E così fa: in 3-4 anni diventa il massimo studioso di topologia, c'è ancora adesso ad esempio un teorema del punto fisso di Brouwer che è l'inizio di un ramo importante della topologia. Dopo di che vince il posto e ricomincia a costruire matematica intuzionistica.

Tra Hilbert e Brouwer c'è un duro scontro di personalità, ed è quello che tutti si divertono a raccontare, anche per parlar male dell'uno o dell'altro. Di solito per parlare male di Brouwer, perché è più facile parlare male di chi è considerato alternativo o in minoranza.

Dallo scontro in poi, praticamente dal punto di vista di "successo numerico" ha vinto Hilbert. Quasi tutta la matematica di oggi è fatta nel modo introdotto da Hilbert sviluppata nel senso introdotto da Hilbert, e la visione costruttiva o intuizionistica della matematica è rimasta minoritaria (in particolare, vi trovate di fronte –per vostra sfortuna?– un rappresentante di questa matematica costruttiva, quindi di questa piccola minoranza).

È ancora minoritaria oggi, ma ha cominciato a suscitare maggior interesse da quando ci si è accorti che è molto rilevante in rapporto ai calcolatori. Infatti, se si vuole raccontare la matematica ad un calcolatore, in modo tale che tutto funzioni bene, questa matematica deve essere scritta usando la logica intuzionistica; perché altrimenti, semplificando al massimo, il programma del calcolatore non contiene gli algoritmi che lo mettono in grado di fornire gli enti di cui dimostra l'esistenza. Immaginate che un giorno ci siano macchine in grado di sviluppare la matematica, di dimostrare teoremi, ecc.; se ad un certo punto una tale macchina dice: "ho dimostrato che la tale equazione ha una soluzione", non vorreste forse anche toccare con mano tale soluzione, vederla? Ebbene, questo sarà pssibile *solo* se le istruzioni date al calcolatore gli permettono di conservare la qualità dell'informazione, e in particolare di mantenere la distinzione tra un vero "esiste" e una doppia negazione come "non è escluso che esista". Cioè, *solo* se la macchina è stata istruita ad usare solo la logica intuizionistica (e una teoria degli insiemi adeguata). È quindi con l'aumento del ruolo dei calcolatori che torna in auge lo studio della matematica costruttiva.

Se ci siamo convinti che ha senso mantenere la distinzione tra logica classica e logica intuizionistica, una bella domanda è: ma allora, per ottenere la matematica costruttiva, si deve rifare tutta la matematica? La risposta è sì, in buona parte. Bisognerebbe ripensare tutta la matematica e vedere quale si può fare e quale no senza usare il principio del terzo escluso o la riduzione all'assurdo.

Forse questo all'inizio rende perplessi, perché si è abituati alla matematica tradizionale, costruita sulla base della logica classica, e si fa fatica ad immaginare quali cambiamenti si dovrebbero fare per far sì che sia basata solo sulla logica intuizionistica. In altre parole, non è facile identificare i punti in cui si usa il principio del terzo escluso, e poi immaginare come si potrebbe farne a meno. E anche non è facile vedere quali vantaggi si possono avere dal non usarlo mai. La risposta è che ci si deve abituare un po' alla volta, provando a fare matematica in pratica. Non c'è una ricetta fissa. Quel che posso dire è che, essendo io stesso un costruttivista, tutto quello che vedremo sulla matematica, negli esempi, sarà perfettamente accettabile anche dal punto di vista costruttivo.

Ma quale è il vantaggio che compensi la fatica di fare la matematica costruttivamente? A parte la questione dei calcolatori, il vantaggio è che la matematica diventa più intuitiva, più facile concettualmente e soprattutto una matematica che conserva la qualità dell'informazione, tanto da poter dire che se si dimostra che esiste qualcosa, si è anche in grado di trovarla.

Dal mio punto di vista –e so di esprimere il punto di vista di una minoranza– si dovrebbe investire molta più energia nello studio della matematica costruttiva. O almeno, si dovrebbe mettere al corrente gli studenti –come sto facendo con voi adesso– del fatto che una matematica costruttiva esiste. Altrimenti gli studenti si familiarizzano con un unico modo di procedere, pensano che esista solo quello e nemmeno si immaginano che esiste un modo alternativo. Con una metafora (introdotta da Brouwer), fare matematica senza mai essere consapevoli della matematica costruttiva sarebbe come insegnare agli studenti di fisica solo la fisica classica, quella di Galileo e Newton, senza mai dire che c'è anche la teoria della relatività di Einstein, o la teoria dei quanti. Per lo meno, io ho cercato di farvi intravedere una possibilità.

Molto di più ci sarebbe da dire su questi temi; a me bastava se non altro indicarvi che esiste una possibilità diversa da quella comune.