Formule strutture teorie

La logica, lo dice il nome, si occupa di linguaggi. Un *linguaggio* formale è un insieme di stringhe, cioè sia sequenze finite, di simboli di un certo *alfabeto*. Da un punto di vista tecnico, se vogliamo formalizzare il nostro discorso, per esempio, nella teoria degli insiemi, potremmo dire che l'alfabeto può essere un insieme qualunque. Se, per esempio, usiamo come alfabeto $A = \{a, b, c \dots z\}$, dove a, b, c, etc. possono essere qualunque purché distinti fra loro, allora le stringhe sono sequenze finite come, per esempio

$$() \qquad (a, x, a, x, a, x, a, s) \qquad (m, l, ö)$$

che conviene scrivere più compattamente nella forma

$$\varepsilon \text{ (stringa vuota)} \qquad axaxaxas \qquad mlö$$

L'ambiente in cui formalizziamo le nostre definizioni, in questo caso la teoria degli insiemi, si dice *metateoria*. Non vogliamo forzare una particolare scelta per la metateoria. Sarà elementare formalizzare il materiale di questo corso, per chi, per esempio, ha seguito il corso di Elementi di Teoria degli Insiemi, prendendo come metateoria la teoria degli insiemi di Zermelo-Fraenkel ZFC. La nostra esposizione sarà

basata su ZFC, mantenendo, però, un tono informale, con due promesse: di non far leva su dettagli incidentali della formalizzazione in ZFC, e di evidenziare i casi in cui si sfruttano principi insiemistici non costruttivi — in pratica, forme dell'assioma di scelta.
In conclusione, parleremo di stringhe che sono sequenze finite di elementi dell'alfabeto, ma mantenendo l'illusione che siano semplicemente tracce di inchiostro sulla carta, che è, poi, l'intuizione che intendiamo formalizzare. — Illusione doppiamente, perché, nella fattispecie, non c'è inchiostro ma configurazioni di elettricità statica.

Formule

Informalmente, vorremmo generalizzare il meccanismo per mezzo del quale si dice che un gruppo è un insieme dotato di un elemento identità e, un'operazione $-\cdot-$, etc. tale che $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, $x \cdot e = e \cdot x = x$, etc. Vorremmo dire che un gruppo è semplicemente un modello della teoria dei gruppi, la quale è costituita dalle condizioni — l'associatività, l'esistenza dell'identità e degli inversi — che deve soddisfare una struttura per essere un gruppo. Per esplorare matematicamente la relazione

che lega una teoria ai suoi modelli, occorre specificare precisamente di che tipo siano le condizioni che possono costituire una teoria — per noi saranno le **formule del primo ordine** — e come cosa sia una struttura che le soddisfa. L'idea è che una formula si ottenga combinando **formule atomiche** — per esempio equazioni — per mezzo di connettivi logici $\wedge \vee \neg \rightarrow$ e i quantificatori $\forall$ e $\exists$. Per esempio:

Es. (Teoria dei gruppi — v. 1)

simboli di base: e, $- \cdot -$, $-^{-1}$

assiomi: $\forall x \forall y \forall z \quad (x \cdot y) \cdot z = x \cdot (y \cdot z)$

$\forall x \quad e \cdot x = x$

$\forall x \quad x \cdot e = x$

$\forall x \quad x \cdot (x^{-1}) = x$

$\forall x \quad x^{-1} \cdot x = x$

Il dominio dei quantificatori è costituito da tutti gli elementi del gruppo: $\forall x$ significa "per ogni elemento x del gruppo".

Es. (Teoria dei gruppi — v. 2)

Spesso non c'è un unico modo di formalizzare un concetto.

simboli di base: e, $- \cdot -$

assiomi: $\forall x \forall y \forall z \quad (x \cdot y) \cdot z = x \cdot (y \cdot z)$

$\forall x \quad e \cdot x = x$

$\forall x \quad x \cdot e = x$

$\forall x \quad \exists y \quad x \cdot y = e \wedge y \cdot x = e$

Crucialmente è ammesso quantificare solo su elementi della struttura, non, per esempio, sui sottoinsiemi della struttura sui numeri naturali, sulle formule stesse. È per questa ragione che la logica che studiamo si dice del primo ordine: se, per esempio, potessimo quantificare anche sui sottoinsiemi della struttura, allora lavoreremmo al secondo ordine. ZFC, la teoria degli insiemi, ricorderete, è formalizzata al primo ordine — c'è un solo tipo di oggetti, gli insiemi, e si può dire "per ogni insieme x" o "esiste un insieme x". Quando, in ZFC, si quantifica sui sottoinsiemi, lo si fa per mezzo di una perifrasi, $\forall x \subseteq y \dots$ significa $\forall x\ x \subseteq y \to \dots$, e questo è sottilmente diverso da dire "per ogni sottoinsieme x di y", infatti, dice "per ogni elemento x dell'universo degli insiemi che sia un sottoinsieme di y". Vedremo una conseguenza sorprendente, il paradosso di Skolem, di questo fatto.

Formalmente, definiremo le formule dando una grammatica — in particolare una grammatica libera dal contesto. Una grammatica identifica le stringhe di un linguaggio descrivendo un processo ricorsivo che permette di scrivere una stringa più lunga combinando stringhe più brevi. Lo studio, in generale, delle grammatiche non fa parte

degli obiettivi di questo corso, vediamo invece il caso particolare che ci interessa.

Def (linguaggio del primo ordine)
Un *linguaggio del primo ordine* — brevemente: *linguaggio* — $L = (R, F, ar)$ è il dato di due insiemi disgiunti R e F di *simboli di relazione* e *simboli di funzione* rispettivamente, e di una funzione $ar : R \cup F \to \mathbb{N}$ che assegna ad ogni simbolo un numero detto *arietà*.

Es. (linguaggio degli anelli ordinati)
Il linguaggio degli anelli ordinati è

$$L_{or} = (\{<\}, \{0, 1, +, \cdot\}, ar_{or})$$

dove

$ar(<) = 2$ — $<$ è un simbolo di relazione binaria

$ar(+) = ar(\cdot) = 2$ — $+$ e $\cdot$ sono simboli di funzione binaria

$ar(0) = ar(1) = 0$ — 0 e 1 sono simboli di costante, che vediamo come funzioni zero-arie.

Nota: qui c'è un piccolo conflitto nella terminologia, perché, secondo la definizione precedente, un "linguaggio" è, in pratica, la collezione dei simboli di base di una teoria, mentre abbiamo già chiamato "linguaggio" un insieme di stringhe. È così, e non è colpa mia.

Nota: nella definizione di linguaggio ammettiamo simboli di funzione zero-ari, che chiameremo simboli di costante e simboli di relazione zero-ari, simboli di costante proposizionale. Le costanti proposizionali ammetteranno due sole interpretazioni: vero e falso.

Per il resto di questo capitolo ~~questa sezione~~ fissiamo un linguaggio del primo ordine $L = (R, F, ar)$.

Def (L-termine)

Gli L-termini sono stringhe nell'alfabeto

$$F \sqcup \{x_0, x_1, x_2 \dots\} \sqcup \{(,), ,\}$$

dove $\sqcup$ indica che gli unendi sono disgiunti.

$$Var \stackrel{def}{=} \{x_0, x_1, x_2 \dots\} = \{x_i \mid i \in \mathbb{N}\}$$

è un insieme numerabile di simboli detti simboli di variabile.

Un L-termine può essere

- un simbolo di variabile $x_i \in Var$
- la stringa $f(t_1, t_2 \dots t_k)$ dove $f \in F$ è un simbolo di funzione, $t_1 \dots t_k$ sono L-termini, e $k = ar(f)$.

Nota: Se c è un simbolo di costante — funzione zero-aria — allora $c()$ è un L-termine. In pratica, omettteremo le parentesi, scrivendo semplicemente c. Similmente useremo, per i simboli che denotano le operazioni aritmetiche,

la comune notazione infissa, per esempio $x_0 + x_1 \cdot x_2$ in luogo di $+(x_0, \cdot(x_1, x_2))$. Infine ci prenderemo la libertà di usare scritture diverse da x_0 x_1 x_2 etc. per i simboli di variabile, es. $x + y \cdot z$, dove non può esserci confusione. Non bisogna confondere le scritture di questo tipo $x + y \cdot z$, che sono abbreviazioni, una stenografia che impieghiamo fra di noi per parlare dei termini, con i termini stessi, che sono gli oggetti definiti formalmente.

Es. Alcuni L_{or}-termini

- $(+(x_0, 1()), x_1)$ $+(+(1(), 1()), 1())$ x_{476}

volgo: $(x + 1) \cdot y$ $(1 + 1) + 1$ $\not\exists$

Def (L-formula)

Le L-formule sono stringhe nell'alfabeto

$$F \sqcup R \sqcup \{x_0, x_1, x_2 \dots\} \sqcup \{(,), ,, \top, \bot, \neg, \wedge, \vee, \to, \forall, \exists\}$$

Una L-formula può essere una formula atomica, ossia:

- $\top$ o $\bot$
- $r(t_1, t_2 \dots t_k)$ dove $r \in R$ è un simbolo di relazione, $t_1 \dots t_k$ sono L-termini, e $k = ar(r)$.
- $t_1 = t_2$ con t_1 e t_2 L-termini

oppure è ottenuta combinando formule atomiche mediante

connettivi logici e quantificatori, ossia:

- $(\neg \varphi)$ con φ L-formula
- $(\varphi \wedge \psi)$
- $(\varphi \vee \psi)$
- $(\varphi \to \psi)$

} con φ e ψ L-formule

- $(\forall x_k \varphi)$
- $(\exists x_k \varphi)$

} con φ L-formula e $x_k \in Var$ simbolo di variabile

La tecnica più immediata per dimostrare un enunciato a proposito di tutte le formule è *l'induzione strutturale* o *sulla complessità delle formule*. Ossia, per dimostrare che tutte le formule godono di una proprietà π, si dimostra che π vale per le formule atomiche, e che π vale per una combinazione a patto che valga per le sue componenti. Similmente si può procedere per induzione sulla complessità dei termini: i casi base sono i simboli di variabile e di costante. La correttezza di questi procedimenti è immediata osservando che si possono giustificare con una semplice induzione aritmetica sulla lunghezza delle formule.

Oss. I casi nelle definizioni di L-formula e L-termine sono disgiunti, ossia, data una L-formula, o un L-termine, questo

oggetto ricade necessariamente in uno e uno solo dei casi della sua definizione.

Questa osservazione ci permette di procedere non solo per induzione, ma anche per ricorsione strutturale, ossia definire una funzione delle formule descrivendo come il valore associato da f a una combinazione dipende da f delle componenti. Vediamo qualche esempio.

Def (sottoformula)
Le sottoformule di una formula φ sono

- φ stessa

inoltre

- se φ è $(\neg\psi)$ anche le sottoformule di ψ
- se φ è $(\psi_1 \wedge \psi_2)$ o $(\psi_1 \vee \psi_2)$ o $(\psi_1 \to \psi_2)$ anche le sottoformule di ψ_1 e ψ_2
- se φ è $(\forall x_k \psi)$ o $(\exists x_k \psi)$ anche le sottoformule di ψ

Def (variabili libere di una formula)
Definiamo prima le variabili di un termine. Dato un L-termine t definiamo $var(t)$ dicendo che

- se $t = x_i \in Var$ $\quad var(t) = \{x_i\}$
- se $t = f(t_1 \dots t_k)$ $\quad var(t) = \bigcup_{i=1}^{k} var(t_k)$

Le variabili libere $vl(\varphi)$ di una formula φ sono definite come segue:

- se φ è $\top$ o $\bot$ $\quad vl(\varphi) = \emptyset$
- se φ è $r(t_1, t_2 \dots t_k)$ $\quad vl(\varphi) = \bigcup_{i=1}^{k} var(t_i)$
- se φ è $t_1 = t_2$ $\quad vl(\varphi) = var(t_1) \cup var(t_2)$

- se φ è $(\neg \psi)$ $\quad vl(\varphi) = vl(\psi)$
- se φ è $(\psi_1 \wedge \psi_2)$ o $(\psi_1 \vee \psi_2)$ o $(\psi_1 \to \psi_2)$ $\quad vl(\varphi) = vl(\psi_1) \cup vl(\psi_2)$
- se φ è $(\forall x_k \psi)$ o $(\exists x_k \psi)$ $\quad vl(\varphi) = vl(\psi) \setminus \{x_k\}$

Oss. L'ultimo caso è cruciale: nella formula $\forall x_k \psi$ e $\exists x_k \psi$ la variabile x_k non è libera, è *legata* dal quantificatore. Si noti che non è richiesto che x_k compaia fra le variabili di ψ.

Es.

$vl((\exists x_1 \; x_0 = x_1 \cdot x_1)) = \{x_0\}$ la variabile x_1 non è libera

$vl(((\exists x_1 \; x_0 = x_1 \cdot x_1) \vee (\exists x_0 \; x_1 = x_0 \cdot x_0))) = \{x_0, x_1\}$

$vl((\forall x_7 \; (\exists x_7 \; x_2 + x_2 = x_4))) = \{x_2, x_4\}$

Es (per chi conosce ZF)

Dimostra dettagliatamente in ZF che la ricorsione strutturale è un procedimento corretto.

Nota: le definizioni date in questa sezione hanno lo

scopo di trasformare gli enunciati matematici in oggetti matematici essi stessi: le formule. Questa sezione serve per descrivere formalmente i nostri oggetti di studio e **non ha un intento normativo**. In particolare, a patto di non causare ambiguità, scriveremo le formule in modo abbreviato in ogni situazione pratica. Così ad esempio

$\forall x \exists y \; y \cdot y \cdot y = x + y$ per $(\forall x_0 (\exists x_1 \; x_1 \cdot x_1 \cdot x_1 = x_0 + x_1))$

$\forall x \; 0 < x \to \exists y \; x = y \cdot y$ per $(\forall x_0 (0 < x_0 \to (\exists x_1 \; x_0 = x_1 \cdot x_1)))$

Cosa significano le formule

Definiremo in questa sezione un concetto di *struttura* che generalizza le familiari strutture algebriche — gruppi, anelli, ordini, etc. Diremo quindi cosa significa che una struttura soddisfa una formula. Questo ci permetterà di precisare la nozione di *conseguenza logica*: ψ è conseguenza di φ se tutte le strutture che soddisfano φ, soddisfano anche ψ. Grazie alla nozione di conseguenza logica, potremo parlare di *teorie* e dei loro *modelli* — per esempio i gruppi saranno i modelli della teoria dei gruppi. Insomma, daremo una *semantica* per la logica del primo ordine, ossia una risposta alla domanda "cosa significano le formule?"

Def (L-struttura)

Fissato il linguaggio $L = (R, F, ar)$, diciamo che una *L-struttura* $M = (D; i)$ è il dato di un insieme D, non vuoto, detto *dominio* della struttura, e di una funzione i, che chiameremo *interpretazione dei simboli*, avente come dominio $R \cup F$ e tale che

$$\forall r \in R \quad i(r) \subseteq D^{ar(r)}$$

$$\forall f \in F \quad i(f) : D^{ar(r)} \to D$$

ossia un simbolo di relazione n-aria è interpretato come un sottoinsieme di D^n, e un simbolo di funzione n-aria come una funzione da D^n a D.

Nota: In molti casi, è chiaro dal contesto quali siano le arietà e le interpretazioni dei simboli di un certo linguaggio. Per esempio, se parliamo della struttura $(\mathbb{Z}; 0, +, -, \cdot, <)$ è chiaro che ci riferiamo alla struttura che ha per dominio $\mathbb{Z}$, nel linguaggio $L = (R, F)$ con $R = \{<\}$, $F = \{0, +, -, \cdot\}$,

$$ar(<) = 2 \quad ar(0) = 0 \quad ar(-) = 1 \quad ar(+) = ar(\cdot) = 2$$

$$i(<) = \{(x, y) \in \mathbb{Z}^2 \mid x < y\}$$

$$i(0) : \{\cdot\} \to \mathbb{Z} \qquad i(0)(\cdot) = 0$$

$$i(-) : \mathbb{Z} \to \mathbb{Z} \qquad i(-)(n) = -n$$

$$i(\overset{+}{\cdot}) : \mathbb{Z} \to \mathbb{Z} \qquad i(\overset{+}{\cdot})(m, n) = m + n$$

Non c'è dubbio che conviene scrivere $(\mathbb{N}; 0, +, \cdot, <)$, e non ci si confonde. Detta M questa struttura, in luogo, per esempio, di $i(+)$, scriveremo semplicemente $+_M$ o anche solo $+$, così $i(+)(2,3) = 2 +_n 3 = 2+3$.

Per il resto di questa sezione fissiamo una L-struttura $M = (D; i)$. Vogliamo ora dire quando una formula φ è *valida* nella struttura M, o, equivalentemente, M *soddisfa* φ, in simboli $M \models \varphi$. Per poter formulare la definizione per ricorsione strutturale, occorre generalizzare il concetto introducendo un *ambiente* o *valutazione delle variabili* v. Così, per esempio $M \models \{v\}\, x = y$ se e solo se l'ambiente v dà a x e y il medesimo valore.

Def (valutazione delle variabili)
Valutazione delle variabili è un modo poetico di dire funzione da Var a D.

Notazione Data una valutazione delle variabili v, e un $a \in D$, indichiamo con $v[^a/_{x_n}]$ la valutazione:

$$v[^a/_{x_n}](x_i) = \begin{cases} v(x_i) & \text{se } i \neq n \\ a & \text{se } i = n \end{cases}$$

Con $v[^{a_1}/_{x_{n_1}} \dots {}^{a_k}/_{x_{n_k}}]$ indichiamo $v[^{a_1}/_{x_{n_1}}] \dots [^{a_k}/_{x_{n_k}}]$

Def (Semantica di Tarski)

Ricordiamo che è fissato un linguaggio L e una L-struttura M. Fissiamo anche un ambiente v. In questo contesto, l'interpretazione di un termine t è definita ricorsivamente da

$$\{v\}_M f(t_1 \dots t_n) \overset{def.}{=} f_M(\{v\}_M t_1 \dots \{v\}_M t_n) \quad \text{con } f \in F$$

$$\{v\}_M x_n \overset{def.}{=} v(x_n)$$

La relazione di soddisfazione $M \models \{v\}\varphi$ per una formula φ nella struttura M e nell'ambiente v è definita ricorsivamente da

$$M \models \{v\} \top \qquad \neg\, M \models \{v\} \bot$$

$$M \models \{v\} r(t_1 \dots t_n) \iff (\{v\}_M t_1 \dots \{v\}_M t_n) \in r_M \quad \text{con } r \in R$$

$$M \models \{v\} t_1 = t_2 \overset{def.}{\iff} \{v\}_M t_1 = \{v\}_M t_2$$

$$M \models \{v\} \neg\psi \overset{def.}{\iff} \neg\, M \models \{v\}\psi$$

$$M \models \{v\} \psi_1 \wedge \psi_2 \overset{def.}{\iff} M \models \{v\}\psi_1 \ \wedge\ M \models \{v\}\psi_2$$

$$M \models \{v\} \psi_1 \vee \psi_2 \overset{def.}{\iff} M \models \{v\}\psi_1 \ \vee\ M \models \{v\}\psi_2$$

$$M \models \{v\} \psi_1 \to \psi_2 \overset{def.}{\iff} M \models \{v\}\psi_1 \ \to\ M \models \{v\}\psi_2$$

$$M \models \{v\} \forall x_n \psi \overset{def.}{\iff} \forall a \in D \quad M \models \{v[a/x_n]\}\psi$$

$$M \models \{v\} \exists x_n \psi \overset{def.}{\iff} \exists a \in D \quad M \models \{v[a/x_n]\}\psi$$

Es. Sia $M = (D; p)$ dove p rappresenta un simbolo di relazione unaria. Cosa significa, secondo la semantica di Tarski, $M \models \{v\} \exists x (p(x) \to \forall y\, p(y))$?