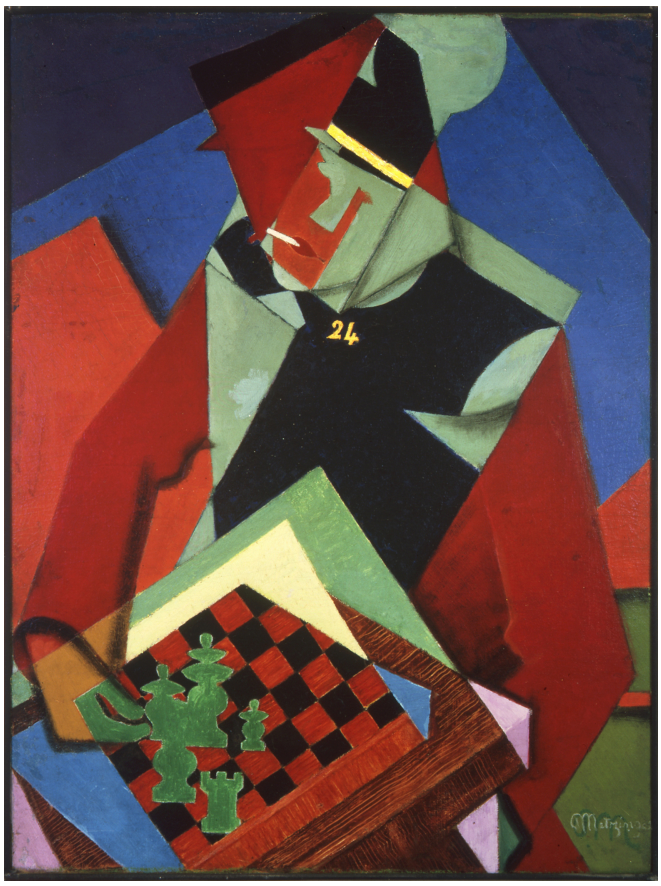
Logica Matematica

APPUNTI DEL CORSO DI LOGICA MATEMATICA
TENUTO DAL PROF. MARCELLO MAMINO

Diego Monaco
d.monaco2@studenti.unipi.it
Università di Pisa

Anno Accademico 2024-25



Indice

Premessa

Queste dispense sono la quasi esatta trascrizione in LaTeX delle dispense del corso di Logica Matematica [1], tenuto dal prof. Marcello Mamino nell'anno accademico 2023-24 presso l'Università di Pisa.

§1 Introduzione

La **logica matematica** nasce dalle ricerche, a cavallo fra il XIX e il XX secolo, sui fondamenti della matematica. Sebbene diverse importanti intuizioni siano considerevolmente più antiche - per esempio il metodo assiomatico degli Elementi di Euclide, l'analisi delle leggi del pensiero nella Logica di Aristotele, l'idea del ragionamento simbolico di Leibniz - si può sostenere che la logica nasca, come branca della matematica, nel momento in cui i concetti quali la **dimostrazione** o la **relazione di conseguenza logica** vengono concepiti come oggetti matematici. In altre parole, la logica diventa matematica nel momento in cui si rende conto che le proposizioni matematiche - scrivibili e scritte in un opportuno linguaggio simbolico - sono passibili di studio alla stessa maniera, per dire, dei polinomi o dei numeri interi. Per cui, per esempio, distinguere quali numeri siano composti, quali polinomi abbiano uno zero, quali proposizioni siano dimostrabili, sono problemi diversi, ma fondamentalmente analoghi.
La logica matematica prende le mosse da un problema spropositato: trasformare l'intera matematica in un gioco di scacchi, e stabilirne - decretarne - le regole. Non è per tracotanza che alcuni pensatori concepirono questo obiettivo, ma per necessità, o quasi per spavento. Infatti, si può tollerare che la fondatezza di metodi come la teoria degli insiemi di Cantor sia argomento - e nutrimento - di speculazione filosofica, sol finché le implicazioni di questa dottrina si limitano all'indagine di un buffo concetto di infinito. Quando matematici di spicco riconoscono la rilevanza dei metodi insiemistici per la matematica nel suo complesso - Hilbert: "Aus dem Paradies, das Cantor uns geschaffen..." - allora è necessario che si giunga ad un consenso sulla correttezza di questi metodi.
Di fronte alla controversia, la matematica reagisce nel solo modo che conosce: occorre capire cosa sia precisamente una dimostrazione, studiare le proprietà di questi oggetti, e dimostrare che non è possibile dimostrare una contraddizione, e, forse, avendo fortuna, dimostrare addirittura che una proposizione può essere dimostrata precisamente quando questo è impossibile per la sua negazione. Ecco, sintetizzato in maniera un po' puerile - ma avete la mia parola che non so fare meglio - il **programma di Hilbert**. Chiaramente, mancano mille dettagli - il più importante: quali metodi sono consentiti nell'esecuzione del piano? Qual è la **metateoria** su cui si deve basare la dimostrazione della solidità delle fondamenta di ogni altra dimostrazione? Possiamo concederci di accantonare questa domanda. Se il programma di Hilbert si potesse portare a termine in una metateoria - una qualunque, che non sia contraddittoria - questa costituirebbe un insieme sufficiente di principi, e si lavorerà poi per scremarli. È chiaro, però, che non si deve barare - se si vuole dimostrare che un sistema assiomatico T è solido, non vale partire da una metateoria MT che ha, fra i suoi assiomi, l'enunciato "T è solido". Intuitivamente, perché l'intera operazione abbia un senso, è meglio che MT sia - a prima vista, almeno - non meno affidabile di T stessa. Tecnicamente, il minimo che si possa pretendere è che MT sia un sottoinsieme di T. In conclusione, il programma di Hilbert richiede, come minimo, di trovare un sistema formale abbastanza vasto da servire come ragionevole fondamento della matematica, e di identificare un segmento di questo sistema che sia intuitivamente valido, e capace di dimostrare la coerenza del sistema nel suo complesso.
Tutti sanno che il programma di Hilbert è deragliato a causa dei **teoremi di incompletezza di Gödel**, del 1931. È andata così: in pratica, un'operazione di hacking. Gödel ha considerato un arbitrario sistema assiomatico T, sotto la condizione che abbia una presentazione effettiva e che sia capace di esprimere una modica quantità di aritmetica - ragionevolmente, qualunque teoria si voglia prendere a fondamento della matematica deve avere queste caratteristiche. Questo sistema T, dimostra Gödel, è in grado di esprimere

proposizioni a proposito di un calcolatore universale, e, in questo calcolatore universale, si può implementare un sistema formale qualunque, per esempio T stesso. Pare, ora, di essere sulla giusta via per il programma di Hilbert: abbiamo T, e dentro T c'è un pezzo di T che è in grado di esprimere proposizioni a proposito delle dimostrazioni di T, questo pezzo vorrà essere MT. Se si vuole seguire il programma di Hilbert, anzi, si deve passare di qua. Qui, però, cominciano i guai. Grazie a un trucco geniale, è possibile sfruttare questa situazione per costruire una proposizione aritmetica che non può né essere dimostrata né confutata in T: una **proposizione indecidibile**. Sfuma quindi la possibilità che T permetta di dirimere ogni possibile questione matematica. Ma c'è di peggio: un'analisi accurata dell'argomento precedente rivela che la proposizione "T non è contraddittoria" è indecidibile in T. A fortiori, quindi, nessuna metateoria che sia un sottoinsieme di T - neppure, appunto, T stessa - può dimostrare la non contraddittorietà di T. Per l'arbitrarietà di T, il programma di Hilbert è rovinato.
Cosa abbiamo imparato da questo disastro? Intanto abbiamo dato una definizione precisa di enunciato o **formula** e una definizione di **deduzione** basata su regole formale, ossia algebriche, simboliche - "abbiamo dato" come comunità matematica, ossia daremo durante il corso. Questa definizione è quella giusta nel senso che, ha dimostrato Gödel nel 1929, c'è una nozione associata di **struttura** - per esempio i gruppi sono precisamente le strutture che soddisfano le formule che esprimono gli assiomi dei gruppi: $\forall x, y, z\ (x \cdot y) \cdot z = x \cdot (y \cdot z)$, $\forall x\ x \cdot e = x$, $\forall x\ e \cdot x = x$, etc. Una formula è una **conseguenza logica** di un certo insieme di formule quando tutte le strutture che soddisfano le formule dell'insieme soddisfano anche la formula. Il **teorema di completezza** di Gödel, del 1929, appunto, garantisce che le nozioni di conseguenza logica e deducibilità coincidono. Fissato il sistema di regole deduttive appena descritto, che chiamiamo **logica del primo ordine** - con riferimento al fatto che è ammesso quantificare $\forall x$, $\exists x$ su elementi della struttura, ma non si può quantificare su suoi sottoinsiemi - i teoremi di incompletezza di Gödel constatano che certe cose non si possono fare. Per esempio non si può dare un'assiomatizzazione effettiva e completa dell'aritmetica. Ce ne faremo una ragione, come ci siamo fatti una ragione del fatto che l'equazione di quinto grado non si può risolvere per radicali o che una primitiva di $\frac{\sin x}{x}$ non si può scrivere come una composizione di funzioni elementari. Resta il fatto che, per arrivare ai risultati di incompletezza di Gödel, è stato necessario costruire, all'interno dell'aritmetica, un calcolatore universale, operazione che certamente involve rendersi conto dell'esistenza di una nozione generale di **funzione calcolabile**, e la costruzione di una **funzione computabile universale** - passi che preludono alla materializzazione elettronica di questi concetti. Va da se che, nel corso, studieremo le basi della **teoria della computabilità**.
Questi argomenti costituiscono quindi l'ossatura tradizione del corso di logica matematica: il calcolo dei predicati del primo ordine, i teoremi di correttezza e completezza del medesimo, alcuni rudimenti di teoria dei modelli, le bassi della teoria della computabilità, e i famosi teoremi di incompletezza di Gödel.
Prima di intraprendere il viaggio, però, è naturale porsi una domanda: non può essere che le limitazioni evidenziate dal fenomeno dell'incompletezza siano, in qualche modo, legate unicamente al particolare sistema di formule e regole deduttive che ci accingiamo a studiare? O forse al metodo assiomatico? Non può darsi che un paio di millenni di abitudine al metodo assiomatico ci abbiano assuefatto all'angustia di questo particolare vicolo cieco, mentre potrebbe esistere un calcolo logico di concezione completamente diversa e immune all'anatema di Gödel, se solo lo cercassimo con mente aperta? No, non c'è via di fuga, ma si può studiare della matematica interessante per capire perché. Intanto, questo è un corollario dell'incompletezza: che l'insieme delle proposizioni aritmetiche vere, espresse nel linguaggio dell'aritmetica del primo ordine, non è computabile. Ossia, non c'è una

funzione computabile che, data in input una proposizione aritmetica, stabilisce se questa sia vera oppure no. Se accettiamo la **tesi di Church**, la quale asserisce che le funzioni computabili in teoria sono praticamente quelle implementabili in pratica, potremmo dire che non è concepibile un programma per computer che distingua le proposizioni aritmetiche vere da quelle false. Questo toglie di mezzo gli assiomi, ma resta il fatto che stiamo parlando di proposizioni scritte nel linguaggio dell'aritmetica del primo ordine. Magari, in questo linguaggio si possono scrivere proposizioni esoteriche e incomprensibili alla matematica ordinaria, proposizioni della cui verità non importa a nessuno. È tutto qui il guaio? Non anche se voi vi credete assolti, a patto che vi importi delle equazioni diofantee, siete coinvolti. È infatti, possibile rafforzare il corollario precedente.

Teorema 1.1 (Davis - Putnam - Robinson (1960) + Matijasevic (1970))
L'insieme dei polinomi $p(x_1, \dots, x_n)$ a coefficienti interi tali che $p(x_1, \dots, x_n) = 0$ abbia soluzione intera, non è calcolabile.

In altri termini, è inconcepibile una procedure - sia essa un sistema assiomatico o qualunque altro tipo di algoritmo - che, ricevendo in input il polinomio p, determina infallibilmente se l'equazione diofantea $p(x_1, \dots, x_n) = 0$ ha soluzione intera. Il decimo problema di Hilbert

Mathematische Probleme.

Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900.

Von

D. Hilbert.

10. Entscheidung der Lösbarkeit einer Diophantischen Gleichung.

Eine Diophantische Gleichung mit irgend welchen Unbekannten und mit ganzen rationalen Zahlencoefficienten sei vorgelegt: *man soll ein Verfahren angeben, nach welchem sich mittelst einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.*

non ha soluzione. La dimostrazione di questo risultato è forse l'apoteosi del metodo che ci ha dato Gödel nel suo lavoro del 1931. Se lì, per dimostrare l'incompletezza dell'aritmetica, si trattava di descrivere un compilatore capace di tradurre ogni funzione computabile in una formula aritmetica, per dimostrare questo teorema occorre compilare ogni funzione computabile in un polinomio.
Non c'è quindi scampo per il programma di Hilbert? Il cardine era l'idea di dimostrare la fondatezza di costruzioni concettuali complesse basandosi su teorie più semplici, e, da ultimo, poggiare tutto sull'aritmetica. Questo cardine è saltato, ma cosa ne è degli obiettivi del programma? Ecco, la situazione ricorda (in qualche senso) quel prigioniero

del mercoledì[1]. Dopo che abbiamo dimostrato oltre ogni ragionevole dubbio la fine del programma di Hilbert, ci guardiamo attorno, e vediamo che esiste una teoria formale su cui i matematici si trovano d'accordo a fondare la matematica: la teoria degli insiemi di Zermelo-Fraenkel espressa nel contesto del calcolo dei predicati al primo ordine. C'è un vasto consenso sulla coerenza di questa teoria. Il livello di precisione formale delle pubblicazioni matematiche è generalmente aumentato nel corso del XX secolo. E chi ha seguito il corso di `Lean 4` sa che ci sono persino diversi matematici di spicco che prendono in seria considerazione la possibilità di formalizzare non solo in teoria, ma in pratica, la matematica per mezzo di opportuni sistemi informatici. Insomma, il cardine è caduto, però il resto del programma non pare che abbia accusato il colpo. Inoltre, fruttuose aree di studio sono nate dalle diramazioni del programma originale: la teoria della dimostrazione, la teoria dei modelli, la teoria degli insiemi, la teoria della computabilità, e forse anche l'informatica teorica. Si può anzi sostenere che il computer, che è nato da molte idee e ha cambiato la faccia dell'umanità, sia, anche, figlio della logica matematica.

[1] In una prigione di qualche remoto paese (Egitto? Texas? Cina?) è domenica e un condannato riceve questa sentenza un po' originale: "Sarai giustiziato prima di domenica prossima, e in nessun modo potrai conoscere il giorno dell'esecuzione fino al giorno stesso ". "Sono salvo!" ragiona il condannato. Le esecuzioni, lo sanno tutti, si fanno solo al mattino. Intanto - lemma 1 - l'esecuzione non può avvenire di sabato, altrimenti, venerdì pomeriggio, il condannato potrebbe conoscere il giorno con cortezza. Quindi venerdì è l'ultimo giorno utile. Per cui - lemma 2 - l'esecuzione non può avvenire venerdì, altrimenti, giovedì pomeriggio etc. Si vede bene che ogni giorno può essere escluso, quindi la sentenza è contraddittoria, e non potrà essere eseguita. Mercoledì mattina, però, il boia (sasin) taglia la testa - o lo avvelena, o fate voi - del condannato, il quale, martedì, non aveva idea del fatto che questa sarebbe stata la sua sorte.

§2 Formule, Strutture e Teorie

La logica, lo dice il nome si occupa di linguaggi. Un **linguaggio** è un insieme di **stringhe**, ossia sequenze finite di simboli, di un certo **alfabeto**[2]. 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, vogliamo usare 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:

$$() \qquad (a, x, a, x, a, x, a, x, a, x) \qquad (m, l, \ddot{o})$$

che conviene scrivere più compattamente come:

$$\varepsilon^{3} \qquad axaxaxaxax \qquad ml\ddot{o}$$

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 fare leva sui dettagli incidentali della formalizzazione di ZFC, e di evidenziare i casi in cui si sfruttano principi insiemistici non costitutivi - 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.

§2.1 Formule

Informalmente, vorremmo generalizzare il meccanismo per mezzo del quale si dice che un **gruppo** è un insieme dotato di un elemento invertibile $\boldsymbol{e}$, un'operazione $\dots \cdot \dots$, 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 vede soddisfare un **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 al primo ordine** - e come una struttura la soddisfa. L'idea è che una formula si ottenga combinando **formule atomiche** - per esempio equazioni - per mezzo di connettivi logici: $\wedge, \vee, \neg, \to$ e i quantificatori $\forall$, $\exists$.

[2] Tecnicamente $L = A^{<\mathbb{N}}$, con A alfabeto.
[3] È la **stringa vuota**.

Esempio 2.1 (Teoria dei gruppi - v.1)

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

$$\forall x\ \forall y\ \forall z\ (x \cdot y) \cdot z = x \cdot (y \cdot z)$$
$$\forall x\ e \cdot x = x$$
$$\forall x\ x \cdot e = x$$
$$\forall x\ x \cdot (x^{-1}) = e$$
$$\forall x\ (x^{-1}) \cdot x = e$$

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

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

Esempio 2.2 (Teoria dei gruppi - v.2)

- **simboli di base**: $e, \dots \cdot \dots$
- **assiomi**:

$$\forall x\ \forall y\ \forall z\ (x \cdot y) \cdot z = x \cdot (y \cdot z)$$
$$\forall x\ e \cdot x = x$$
$$\forall x\ x \cdot e = x$$
$$\forall x\ x \cdot (x^{-1}) = e$$
$$\forall x\ \exists y\ x \cdot y = e \wedge y \cdot x = e$$

Crucialmente è ammesso quantificare solo su elementi della struttura, non, per esempio, sui sottoinsiemi. Questo è il motivo per cui si parla di **logica al primo ordine**. della struttura, sui numeri naturali o 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, dire "per ogni elemento 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.

Definizione 2.3 (Linguaggio del primo ordine)**.** Un **linguaggio del primo ordine** - brevemente linguaggio - $L = (R, F, \mathrm{ar})$ è dato da due insiemi disgiunti R e F, rispettiva-

mente i **simboli di relazione** e i **simboli di funzione**, e una funzione $\mathrm{ar} : R \sqcup F \to \mathbb{N}$ che associa ad ogni simbolo un numero naturale, detto **arietà**.[4]

Esempio 2.4 (Linguaggio degli anelli ordinati)
Il linguaggio degli anelli ordinati è:

$$L_{or} = (\{<\}, \{0, 1, +, \cdot\}, \mathrm{ar}_{or})$$

dove:

$\mathrm{ar}_{or}(<) = 2$	$<$ è un simbolo di relazione binaria
$\mathrm{ar}_{or}(+) = \mathrm{ar}_{or}(\cdot) = 2$	$+$ e $\cdot$ sono simboli di funzione binaria
$\mathrm{ar}_{or}(0) = \mathrm{ar}_{or}(1) = 0$	0 e 1 sono simboli di costante

Si osservi che i simboli di constante li vediamo come funzioni di arietà 0.

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" l'insieme delle stringhe. È così non è colpa mia.

Osservazione 2.5 — Nella definizione di linguaggio ammettiamo simboli di funzione 0-ari, che chiameremo **simboli di costante**, e simboli di relazione 0-ari, che chiameremo **simboli di costante proposizionale**. Le costanti proposizionali ammetteranno due sole interpretazioni: *vero* e *falso*.

Per il resto di questo capitolo fissiamo un linguaggio al primo ordine $L = (R, F, \mathrm{ar})$.

Definizione 2.6 (L-termine)**.** Gli **L-termini** sono stringhe dell'alfabeto dato da:

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

Chiamiamo l'insieme numerabile:

$$\mathrm{Var} \stackrel{\text{def}}{=} \{x_0, x_1, x_2, \ldots\} = \{x_i\}_{i \in \mathbb{N}}$$

insieme dei **simboli di variabile**. Un **L-termine** è quindi una stringa in $F \sqcup \mathrm{Var} \sqcup \{(,), ,\}$, e può essere [definito induttivamente come segue]:

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

Osservazione 2.7 — Se c è un simbolo di costante - funzione 0-aria - allora $c()$ è un L-termine (abbiamo detto che le funzioni k-arie valutate in L-termini sono a loro volta L-termini, dunque $c()$ lo è in automatico). In pratica, ometteremo 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.

[4] Tecnicamente staremmo anche fissando un alfabeto da cui prendere i simboli.

Non bisogna confondere le scritture di questo tipo $x + y \cdot z$, che sono abbreviazioni, un stereografia che impieghiamo fra di noi per parlare dei termini, con i termini stessi, che sono gli oggetti definiti formalmente.

Esempio 2.8
Ecco alcuni esempi di L_{or} termini:

$$\cdot(+(x_0, 1()), x_1) \qquad +(+(1(), 1()), 1())$$

vulgo:

$$(x_0 + 1) \cdot x_1 \qquad (1 + 1) + 1$$

Definizione 2.9 (L-formula)**.** Le **L-formule** sono stringhe dell'alfabeto dato da:

$$F \sqcup R \sqcup \mathrm{Var} \sqcup \{(,), ,, \top, \bot, \neg, \wedge, \vee, \to, \forall, \exists\}$$ [5]

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

- $\top$ o $\bot$,
- $r(t_1,t_2,\dots,t_k)$ con $r \in R$ simbolo di relazione e $t_1, \dots, t_k$ L-termini, e $\mathrm{ar}(r) = k$,
- $t_1 = t_2$ con t_1, t_2 L-termini.

oppure è ottenuta combinando formule atomiche per mezzo di **connettivi logici** e **quantificatori**:

- $(\neg\, \varphi)$ con φ L-formula,
- $(\varphi \wedge \psi)$ con φ, ψ L-formule,
- $(\varphi \vee \psi)$ con φ, ψ L-formule,
- $(\varphi \to \psi)$ con φ, ψ L-formule,
- $(\forall x_k\, \varphi)$ con φ L-formula e $x_k \in \mathrm{Var}$ simbolo di variabile,
- $(\exists x_k\, \varphi)$ con φ L-formula e $x_k \in \mathrm{Var}$ simbolo di variabile.

La tecnica più immediata per dimostrare un enunciato a proposito di tutte le formule è l'**induzione strutturale** o **induzione 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.

[5] Pertanto la differenza sostanziale tra L-termini ed L-formule sta nel fatto che, nelle seconde, le stringe possono essere costruite ricorsivamente anche usando connettivi logici e quantificatori (ed usando come base anche relazioni di L-termini (e non funzioni)).

Osservazione 2.10 — 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 un 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 solo da f e dalle sue componenti. Vediamo qualche esempio.

Definizione 2.11 (Sottoformula)**.** Le **sottoformule** di una formula φ sono:

- φ stessa

inoltre:

- se $\varphi = (\neg\psi)$, allora tra le sottoformule di φ ci sono anche le sottoformule di ψ,
- se $\varphi = (\psi_1 \wedge \psi_2)$ o $\varphi = (\psi_1 \vee \psi_2)$ o $\varphi = (\psi_1 \to \psi_2)$, allora tra le sottoformule di φ ci sono anche le sottoformule di ψ_1 e ψ_2,
- se $\varphi = (\forall x_k\,\psi)$ o $\varphi = (\exists x_k\,\psi)$, allora tra le sottoformule di φ ci sono anche le sottoformule di ψ.

Nella definizione precedente abbiamo definito le sottoformule per via ricorsiva - ossia, abbiamo definito le sottoformule di una formula φ in termini delle sottoformule delle componenti di φ (oltre a φ stessa).

Definizione 2.12 (Variabili libere di una formula)**.** Definiamo prima le **variabili libere di un termine**. Dato un L-termine t definiamo $\mathrm{var}(t)$ dicendo che:

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

Adesso che abbiamo definito ricorsivamente le variabili libere di un L-termine t, possiamo definire le **variabili libere di una L-formula** φ, $\mathrm{vl}(\varphi)$ come segue per le formule atomiche:

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

e per le formule composte:

- se $\varphi = (\neg\psi)$, allora $\mathrm{vl}(\varphi) = \mathrm{vl}(\psi)$,
- se $\varphi = (\psi_1 \wedge \psi_2)$ o $\varphi = (\psi_1 \vee \psi_2)$ o $\varphi = (\psi_1 \to \psi_2)$, allora $\mathrm{vl}(\varphi) = \mathrm{vl}(\psi_1) \cup \mathrm{vl}(\psi_2)$,
- se $\varphi = (\forall x_k\,\psi)$ o $\varphi = (\exists x_k\,\psi)$, allora $\mathrm{vl}(\varphi) = \mathrm{vl}(\psi) \setminus \{x_k\}$.

Osservazione 2.13 (Variabili legate) **—** L'ultimo caso è cruciale: nelle formule $\forall x_k\,\psi$ 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 libere di ϕ [Possibile typo di Mamino ψ?].

Esempio 2.14

Vediamo alcuni esempi:

- ♢ $\mathrm{vl}(\exists x_1\ x_0 = x_1 \cdot x_1) = \{x_0\}$ e la variabile x_1 è legata.
- ♢ $\mathrm{vl}((\exists x_1\ x_0 = x_1 \cdot x_1) \wedge (\exists x_0\ x_1 = x_0 \cdot x_0) = \{x_0\} \cup \{x_1\} = \{x_0, x_1\}$[a].
- ♢ $\mathrm{vl}((\forall x_7\ (\exists x_7\ x_2 + x_2 = x_4))) = \{x_2, x_4\}$, si noti che il $\forall$ all'inizio è uno specchietto per le allodole, in quanto non conta nulla per il significato della formula.

[a] Se interpretassimo queste due affermazioni nei naturali, la prima vorrebbe dire ogni numero è un quadrato, mentre la seconda che ogni numero naturale ha un quadrato.

Esercizio 2.15 (Per chi conosce ZF)**.** Dimostrare dettagliatamente in ZF che la ricorsione strutturale è un procedimento corretto.

Nota 2.16 — 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[a]. In particolare, a patto di non causare ambiguità, scriviamo le formule in modo abbreviato in ogni situazione pratica. Così ad esempio:

$\forall x\ \exists y\ y \cdot y \cdot y = x + y$	al posto di $\forall x_0(\exists x_1\ x_1 \cdot x_1 \cdot x_1 = x_0 + x_1)$
$\forall x\ 0 < x \to \exists y\ x = y \cdot y$	al posto di $\forall x_0(0 < x_0 \to \exists x_1\ x_0 = x_1 \cdot x_1)$

[a] Non stiamo dicendo come la matematica dovrebbe essere fatta/scritta, stiamo constatando che viene fatta/scritta in certi modi, e stiamo formalizzando precisamente cosa sono questi modi.

§2.2 Cosa significano le formule

Definiamo 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 logica di φ se ogni struttura che soddisfa φ soddisfa 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 lla domanda "cosa significano le formule?".

Definizione 2.17 (L-struttura)**.** Fissato un linguaggio al primo ordine $L = (R, F, \mathrm{ar})$, 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 $F \sqcup R$ (il nostro alfabeto) e tale che:

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

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 2.18 — 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 $R = \{<\}$ e $F = \{0, +, -, \cdot\}$, con:

$$\operatorname{ar}(<) = 2 \qquad \operatorname{ar}(+) = \operatorname{ar}(\cdot) = 2 \qquad \operatorname{ar}(-) = 1 \qquad \operatorname{ar}(0) = 0$$

con:

$$\begin{aligned}
&i(<) = \{(x, y) \in \mathbb{Z}^2 \mid x < y\} \subseteq \mathbb{Z}^2 = D^{\operatorname{ar}(<)} \\
&i(0) : D^{\operatorname{ar}(0)} = \{\bullet\} \in \mathbb{Z} = D \qquad i(0)(\bullet) = 0 \\
&i(-) : D^{\operatorname{ar}(-)} = \mathbb{Z} \to \mathbb{Z} \qquad i(-)(n) = -n \\
&i(+) : D^{\operatorname{ar}(+)} = \mathbb{Z}^2 \to \mathbb{Z} \qquad i(+)(n, m) = n + m \\
&i(\cdot) : D^{\operatorname{ar}(\cdot)} = \mathbb{Z}^2 \to \mathbb{Z} \qquad i(\cdot)(n, m) = n \cdot m
\end{aligned}$$

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

Per il resto di questa sezione fissiamo una L-struttura $M = (D; i)$. Vogliamo 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.

Definizione 2.19 (Valutazione delle variabili)**.** **Valutazione delle variabili** è un modo poetico per dire funzione da Var a D.

Notazione 2.20 — 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 $v[a_1/x_{n_1}] \dots [a_k/x_{n_k}]$.

Definizione 2.21 (Semantica di Tarski)**.** Ricordiamo che è fissato un linguaggio al primo ordine L ed una L-struttura M. Fissiamo anche un ambiente v. In questo contesto possiamo definire per ricorsione strutturale l'**interpretazione** di una L-formula, partendo dagli L-termini come segue:

$$\begin{aligned}
\{v\}_M\, x_k &\stackrel{\text{def}}{=} v(x_k) \\
\{v\}_M\, f(t_1, \dots, t_k) &\stackrel{\text{def}}{=} i(f)\,(\{v\}_M\, t_1, \dots, \{v\}_M\, t_k) \qquad \text{con } f \in F
\end{aligned}$$

La relazione di **soddisfacibilità** $M \models \{v\}\varphi$ (ometteremo M al pedice d'ora in poi) per una formula φ nella struttura M e nell'ambiente v è definita ricorsivamente per le formule

atomiche da:

$$M \models \{v\}\top \qquad \neg M \models \{v\}\bot$$
$$M \models \{v\}\, r(t_1,\dots,t_k) \overset{\text{def}}{\iff} (\{v\}_M\, t_1,\dots,\{v\}_M\, t_k) \in i(r) =: r_M$$
$$M \models \{v\}\, t_1 = t_2 \overset{\text{def}}{\iff} \{v\}_M\, t_1 = \{v\}_M\, t_2$$

ed infine la soddisfacibilità per le L-formule composte è definita come segue:

$$\begin{aligned}
M \models \{v\}(\neg\psi) &\overset{\text{def}}{\iff} \neg M \models \{v\}\psi \\
M \models \{v\}(\psi_1 \wedge \psi_2) &\overset{\text{def}}{\iff} M \models \{v\}\psi_1 \wedge M \models \{v\}\psi_2 \\
M \models \{v\}(\psi_1 \vee \psi_2) &\overset{\text{def}}{\iff} M \models \{v\}\psi_1 \vee M \models \{v\}\psi_2 \\
M \models \{v\}(\psi_1 \to \psi_2) &\overset{\text{def}}{\iff} M \models \{v\}\psi_1 \to M \models \{v\}\psi_2 \\
M \models \{v\}(\forall x_k\, \psi) &\overset{\text{def}}{\iff} \forall a \in D \quad M \models \{v[a/x_k]\}\psi \\
M \models \{v\}(\exists x_k\, \psi) &\overset{\text{def}}{\iff} \exists a \in D \quad M \models \{v[a/x_k]\}\psi
\end{aligned}$$

Esempio 2.22
Sia $(M; p)$, dove p è un simbolo di relazione unaria. Cosa significa, secondo la semantica di Tarski, che $M \models \{v\}\, \exists x(p(x) \to \forall y\, p(y))$?

Soluzione. Intuitivamente, ci aspettiamo che asserire che M soddisfa quella formula equivalga, nella metateoria, alla proposizione $\exists a \in D(a \in p_M \to \forall b \in D\ b \in p_M)$. Vediamo come questo segue formalmente dalla semantica di Tarski.

$$\begin{aligned}
&M \models \{v\}(\exists x(p(x) \to \forall y\, p(y)) \\
&\exists a \in D\ M \models \{v[a/x]\}(p(x) \to \forall y\, p(y)) \\
&\exists a \in D\ M \models \{v[a/x]\}p(x) \to M \models \{v[a/x]\}\forall y\, p(y) \\
&\exists a \in D\ \{v[a/x]\}_M p(x) \to \forall b \in D\ M \models \{v[a/x, b/y]\}p(y) \\
&\exists a \in D\ \{v[a/x]\}_M p(x) \to \forall b \in D\ \{v[a/x, b/y]\}_M p(y) \\
&\exists a \in D\ a \in p_M \to \forall b \in D\ b \in p_M
\end{aligned}$$

□

§3 Eliminazione dei quantificatori

Definizione 3.1 (Formule equivalenti per una teoria)**.** Siano φ e ψ L-formule e T una L-teoria. Diciamo che φ è **equivalenti** per T, denotato con $T \models \varphi \leftrightarrow \psi$ se $T \models (\varphi \to \psi) \wedge (\psi \to \varphi)$, ossia $T, \varphi \models \psi$ e $T, \psi \models \varphi$.

Definizione 3.2 (Eliminazione dei quantificatori)**.** La L-teoria T ha l'**eliminazione dei quantificatori** se ogni L-formula è equivalente, per T, ad una formula senza quantificatori.

Se una teoria coerente T ha l'eliminazione dei quantificatori, e, per ogni formula chiusa e senza quantificatori φ, vale $T \models \varphi$ o $T \models \neg\varphi$, allora T è completa.
Dimostreremo la seguente proposizione.

Proposizione 3.3
La teoria degli ordini lineari densi senza estremi T_{oldse} ha l'eliminazione dei quantificatori.

Corollario 3.4
T_{oldse} è completa.

Vediamo la dimostrazione del corollario.

Dimostrazione. T_{oldse} è coerente perché $\mathbb{Q} \models T_{oldse}$. Inoltre, nel linguaggio di T_{oldse}, non ci sono simboli di costante, quindi le uniche formule chiuse senza quantificatori sono combinazioni booleane di $\top$ e $\bot$. È immediato che queste formule hanno un valore di verità definito. □

Altre teorie con l'eliminazione dei quantificatori sono, per esempio, $\mathrm{Th}(\mathbb{C}, 0, 1, +, \cdot)$ e $\mathrm{Th}(\mathbb{R}, 0, 1, +, \cdot, <)$: rispettivamente, la teoria dei campi algebricamente chiusi e la teoria dei campi reali chiusi.

Osservazione 3.5 — Se φ_1 è una sottoformula di ψ_1, e rimpiazziamo una occorrenza di φ_1 in ψ_1 con una φ_2, che soddisfa $T \models \varphi_1 \leftrightarrow \varphi_2$, allora la formula ψ_2 ottenuta da questa sostituzione soddisfa $T \models \psi_1 \leftrightarrow \psi_2$.

Dell'osservazione. Induzione strutturale. □

Introduciamo due nozioni di logica proposizionali che sono elementari ma spesso utili.

Definizione 3.6 (CNF e DNF)**.** Una formula senza quantificatori φ è in **forma normale congiuntiva (CNF)** se:

$$\varphi = (\alpha_{11} \vee \alpha_{12} \vee \ldots \vee \alpha_{1n_1}) \wedge \ldots \wedge (\alpha_{m1} \vee \alpha_{m2} \vee \ldots \vee \alpha_{mn_m})$$

ovvero φ una congiunzione di disgiunzioni di formule α_{ij} che possono essere formule atomiche o negazioni di formule atomiche. Simmetricamente, φ è in **forma normale disgiuntiva (DNF)** se:

$$\varphi = (\alpha_{11} \wedge \alpha_{12} \wedge \ldots \wedge \alpha_{1n_1}) \vee \ldots \vee (\alpha_{m1} \wedge \alpha_{m2} \wedge \ldots \wedge \alpha_{mn_m})$$

ovvero φ una disgiunzione di congiunzioni di formule α_{ij} che possono essere formule atomiche o negazioni di formule atomiche.

Lemma 3.7
Data una formula φ senza quantificatori, esistono ψ_1 in CNF e ψ_2 in DNF equivalenti a φ per la teoria vuota (e quindi per ogni teoria).

Esercizio. □

T_{oldse} ha l'eliminazione dei quantificatori. Per induzione strutturale, e sfruttando il fatto che $\models (\forall x_k\, \varphi) \to (\neg\exists x_k \neg \varphi)$, vediamo che ci bastano, data una formula φ senza quantificatori, trovare ψ senza quantificatori tale che $T_{oldse} \models (\exists x_k\, \varphi) \leftrightarrow \psi$. □

Esercizio 3.8. Da mettere.

§3.1 Ultraprodotti

Introduciamo una prima tecnica per costruire modelli. Usando gli ultraprodotti dimostreremo il seguente.

Teorema 3.9 (Compattezza - versione semantica)
Data una L-teoria T e una L-formula φ, se $T \models \varphi$ allora esiste un sottoinsieme finito $T' \subseteq T$, tale che $T' \models \varphi$.

Ovvero se φ è conseguenza logica di un insieme di premesse T, allora basta, in realtà, una quantità finita di queste premesse per implicare φ. Poco da stupirsi se si pensa alla conseguenza logica come dimostrabilità: infatti una dimostrazione - non abbiamo ancora formalizzato questo concetto - è un argomento di lunghezza finita, non ha quindi spazio per riferirsi a più di una quantità finita di premesse. D'altro canto, sia l'enunciato sia la dimostrazione che vedremo sono puramente semantici: la nozione di conseguenza logica che stiamo considerando è verità in tutte le strutture che soddisfano le premesse.

Per mostrare un'applicazione squisitamente matematica degli ultraprodotti, dimostreremo altresì il seguente.

Teorema 3.10 (Ax-Grothendieck)
Sia $f : \mathbb{C}^n \to \mathbb{C}^n$ una funzione polinomiale iniettiva, allora f è surgettiva.

Bando alle ciance.

Definizione 3.11 (Filtri ed ultrafiltri)**.** Sia I un insieme fissato, un **filtro su $\mathcal{P}(I)$** è un sottoinsieme $F \subseteq \mathcal{P}(I)$ tale che:

(i) $\emptyset \notin F$ e $I \in F$;

(ii) $A \in F \wedge A \subseteq B \implies B \in F$;

(iii) se $A, B \in F$ allora $A \cap B \in F$.

Un filtro U su $\mathcal{P}(I)$ è un **ultrafiltro** se:

$$A \in U \vee I \setminus A \in U$$

Esempio 3.12
Fissato $I \neq \emptyset$ e $x \in I$, l'insieme:

$$U_x = \{A \subseteq I \mid x \in A\}$$

è un ultrafiltro, detto **ultrafiltro principale generato da x**. Nonostante il nome altisonante, questi ultrafiltri qui servono a poco. Noi abbiamo bisogno di ultrafiltri non principali.

Esempio 3.13
Sia I un insieme infinito. L'insieme:

$$F = \{A \subseteq I : |I \setminus A| < \aleph_0\}$$

dei sottoinsiemi **cofiniti** di I è un filtro su $\mathcal{P}(I)$.

Proposizione 3.14 (Esistenza degli ultrafiltri)
Se F è un filtro su $\mathcal{P}(I)$ allora esiste un ultrafiltro U su $\mathcal{P}(I)$ tale che $F \subseteq U$.

Dimostrazione. Zorn. □

Esempio 3.15
Sia I un insieme infinito, allora esiste un ultrafiltro non principale su $\mathcal{P}(I)$. Basta infatti considerare un ultrafiltro U che estende il filtro dei cofiniti. Se U fosse principale, allora esisterebbe un $x \in I$ avremmo $\{x\} \in U$, ma anche $I \setminus \{x\} \in U$ perché cofinito, da cui $\emptyset = \{x\} \cap (I \setminus \{x\}) \in U$ ↯.

Definizione 3.16 (Ultraprodotto)**.** Fissiamo un linguaggio $L = (R, F)$ e una famiglia di L-strutture $\{M_i\}_{i \in I}$. Sia U un ultrafiltro su I e definiamo l'**ultraprodotto** $\prod_{i \in I} {}^{M_i}\!/_U$, la L-struttura che ha come dominio $\prod_{i \in I} D_i$ modulo la relazione di equivalenza:

$$\forall a, b \in \prod_{i \in I} D_i \quad a \sim b \iff \{i \in I \mid a_i = b_i\} \in U$$

L'interpretazione di $r \in R$ in questo dominio è:

$$([a_1], \dots, [a_k]) \in {}^{r}\!/_U \overset{\text{def}}{\iff} \{i \in I \mid (a_{1i}, \dots, a_{ki}) \in r^{M_i}\} \in U$$

L'interpretazione di $f \in F$ è:

$${}^{f}\!/_U([a_1], \dots, [a_k]) \overset{\text{def}}{=} [\{f_{M_i}(a_{1i}, \dots, a_{ki})\}_{i \in I}]$$

Osservazione 3.17 — Occorre verificare che la definizione è ben posta, ossia che se $a_1 \sim_U a_1', \dots a_k \sim_U a_k'$ allora:

$$([a_1], \dots, [a_k]) \in {}^{r}\!/_U \iff ([a_1'], \dots, [a_k']) \in {}^{r}\!/_U$$

ovvero:

$$f/_U([a_1],\dots,[a_k]) \sim_U f/_U([a'_1],\dots,[a'_k])$$

Riferimenti bibliografici

[1] Marcello Mamino, *Logica Matematica*, Università di Pisa, Pisa, 2023-24.

[2] Marcello Mamino, *Logica Matematica*, Università di Pisa, Pisa, 2024-25.