

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

1 Introduzione	4
2 Formule, Strutture e Teorie	8
2.1 Formule	8
2.2 Cosa significano le formule	13
2.3 Sostituzioni	17
2.4 Teorie	20
2.5 Esempi di teorie	21
2.6 PA e Q di Robinson	23
3 Eliminazione dei quantificatori, forme normali ed ultrafiltri	25
3.1 Eliminazione dei quantificatori	25
3.2 CNF e DNF	27
3.3 Ultraprodotti	29
3.4 Teorema di Ax-Grothendieck	33
3.5 Compattezza semantica e teorema di Loś	35
3.6 Applicazioni del teorema di compattezza	37
3.7 Teoremi di Löwenheim-Skolem	45
3.8 Categoricità e completezza	46
4 Sintassi	48
4.1 Sistema di deduzione naturale (ND)	48
4.2 Sistema ridotto ($ND_{\rightarrow, \perp, \exists, =}$)	48
4.3 Teoremi di correttezza e completezza	48
Bibliografia	50

Premessa

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

Ringraziamenti

Antonio De Lucreziis.

Quest'opera è stata rilasciata con la licenza Creative Commons Attribuzione - Condividi allo stesso modo 4.0 Internazionale. Per leggere una copia della licenza visita il sito web <https://creativecommons.org/licenses/by-nc/4.0/deed.it>.



§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 sé 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 procedura - 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 mercoledì¹. 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.

¹In 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 certezza. 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:

$$() \quad (a, x, a, x, a, x, a, x, a, x) \quad (m, l, ö)$$

che conviene scrivere più compattamente come:

$$\varepsilon^3 \quad axaxaxaxax \quad mlö$$

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 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 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: $\wedge, \vee, \neg, \rightarrow$ e i quantificatori \forall, \exists .

²Tecnicamente $L = A^{<\mathbb{N}}$, con A alfabeto.

³È la **stringa vuota**.

Esempio 2.1 (Teoria dei gruppi - v.1)

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

- **assiomi:**

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

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, \dots

- **assiomi:**

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

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 $\forall x x \subseteq y \rightarrow \dots$, 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 (CFG)**.

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 - $L = (R, F, \text{ar})$ è dato da due insiemi disgiunti R e F , rispettivamente

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

Esempio 2.4 (Linguaggio degli anelli ordinati)

Il linguaggio degli anelli ordinati è:

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

dove:

$$\text{ar}_{or}(\langle) = 2$$

\langle è un simbolo di relazione binaria

$$\text{ar}_{or}(+) = \text{ar}_{or}(\cdot) = 2$$

$+$ e \cdot sono simboli di funzione binaria

$$\text{ar}_{or}(0) = \text{ar}_{or}(1) = 0$$

0 e 1 sono simboli di costante

Si osservi che i simboli di costante 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 $L = (R, F, \text{ar})$.

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

$$F \sqcup \{\textcolor{red}{x}_0, \textcolor{red}{x}_1, \textcolor{red}{x}_2, \dots\} \sqcup \{(), ,\}$$

Chiamiamo l'insieme numerabile:

$$\text{Var} \stackrel{\text{def}}{=} \{\textcolor{red}{x}_0, \textcolor{red}{x}_1, \textcolor{red}{x}_2, \dots\} = \{\textcolor{red}{x}_i\}_{i \in \mathbb{N}}$$

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

- un simbolo di variabile $\textcolor{red}{x}_i \in \text{Var}$
- la stringa $\textcolor{red}{f}(t_1, t_2, \dots, t_k)$, dove $\textcolor{red}{f} \in F$ è un simbolo di funzione, t_1, \dots, t_k sono *L-termini*, e $\text{ar}(f) = k$.

Moralmente: si parte dai simboli di variabile e si applicano simboli di funzione ricorsivamente per costruire termini più complessi.

Osservazione 2.7 (I simboli di costante sono in automatico L-termini) — Se $\textcolor{red}{c}$ è un simbolo di costante - funzione 0-aria - allora $\textcolor{red}{c}()$ 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 $\textcolor{red}{c}$. Similmente useremo, per i simboli che denotano le operazioni aritmetiche, la comune

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

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, un stereografia che impieghiamo fra di noi per parlare dei termini, con i termini stessi, che sono gli oggetti definiti formalmente.

Esempio 2.8 (L -termini)

Ecco alcuni esempi di L_{or} termini:

$$\cdot(+(\mathbf{x}_0, 1()), \mathbf{x}_1) \quad + (+(\mathbf{1}(), 1()), 1())$$

vulgo:

$$(\mathbf{x}_0 + 1) \cdot \mathbf{x}_1 \quad (1 + 1) + 1$$

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

$$F \sqcup R \sqcup \text{Var} \sqcup \{(,), \cdot, \top, \perp, \neg, \wedge, \vee, \rightarrow, \forall, \exists\}^5$$

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

- \top o \perp ,
- $r(t_1, t_2, \dots, t_k)$ con $r \in R$ simbolo di relazione e t_1, \dots, t_k L -termini, e $\text{ar}(r) = k$,
- $t_1 = t_2$ 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,
- $(\varphi \rightarrow \psi)$ con φ, ψ L -formule,
- $(\forall x_k \varphi)$ con φ L -formula e $x_k \in \text{var}$ simbolo di variabile,
- $(\exists x_k \varphi)$ con φ L -formula e $x_k \in \text{var}$ 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.

⁵Pertanto la differenza sostanziale tra L -termini ed L -formule sta nel fatto che, nelle seconde, le stringhe 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 $\varphi = (\psi_1 \rightarrow \psi_2)$, 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 $\text{var}(t)$ dicendo che:

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

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

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

e per le formule composte:

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

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 ψ ^a.

^aPossibile typo di Mamino(?)

Esempio 2.14

Vediamo alcuni esempi:

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

^aSe 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:

$$\begin{aligned} \forall x \exists y y \cdot y \cdot y = x + y & \quad \text{al posto di } \forall x_0 (\exists x_1 x_1 \cdot x_1 = x_0 + x_1) \\ \forall x 0 < x \rightarrow \exists y x = y \cdot y & \quad \text{al posto di } \forall x_0 (0 < x_0 \rightarrow \exists x_1 x_0 = x_1 \cdot x_1) \end{aligned}$$

^aNon 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 alla domanda “cosa significano le formule?”.

Definizione 2.17 (*L*-struttura). Fissato un linguaggio al primo ordine $L = (R, F, \text{ar})$, 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:

$$\begin{aligned} \forall r \in R \quad i(r) &\subseteq D^{\text{ar}(r)} \\ \forall f \in F \quad i(f) : D^{\text{ar}(f)} &\rightarrow D \end{aligned}$$

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 $(\mathbb{Z}, \mathbf{0}, +, -, \cdot, <)$ è chiaro che ci riferiamo alla struttura che ha per dominio \mathbb{Z} , nel linguaggio $L = (R, F)$ con $R = \{<\}$ e $F = \{\mathbf{0}, +, -, \cdot\}$, con:

$$\text{ar}(<) = 2 \quad \text{ar}(+) = \text{ar}(\cdot) = 2 \quad \text{ar}(-) = 1 \quad \text{ar}(\mathbf{0}) = 0$$

con:

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

Non c'è dubbio che conviene scrivere $(\mathbb{Z}; \mathbf{0}, +, -, \cdot, <)$, 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 $v[a/x_n] : \text{Var} \rightarrow D$ 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 e definiamo i concetti di interpretazione di un L -termine e di soddisfabilità di una L -formula. L'**interpretazione** degli L -termini come segue:

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

Invece la relazione di **soddisfabilità** $M \models \{v\}\varphi$ (ometteremo M al pedice d'ora in poi) per una L -formula φ nella struttura M e nell'ambiente v è definita ricorsivamente, a partire dalle formule atomiche, come segue:

$$M \models \{v\} \top \quad \neg M \models \{v\} \perp$$

$$\begin{aligned} M \models \{v\} r(t_1, \dots, t_k) &\stackrel{\text{def}}{\iff} (\{v\}_M t_1, \dots, \{v\}_M t_k) \in i(r) =: r_M \\ M \models \{v\} t_1 = t_2 &\stackrel{\text{def}}{\iff} \{v\}_M t_1 = \{v\}_M t_2 \end{aligned}$$

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

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

Esempio 2.22

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

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

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

□

Incidentalmente, possiamo notare che abbiamo ottenuto una proposizione che non dipende dall'ambiente v . Questo accade perché la formula data è **chiusa**, ossia non ha variabili libere. Inoltre, grazie al fatto che D è non vuoto, la proposizione $\exists a \in D(a \in p_M \rightarrow \forall b \in D b \in p_M)$ è necessariamente vera indipendentemente da M . Infatti si danno due casi, o la conseguente è sempre vera, per cui l'implicazione è sempre vera indipendentemente dall' $a \in D$ che usiamo (qui serve $D \neq \emptyset$), o esiste un $b \in D$ tale che $b \notin p_M$, ma scegliendo a come quel b otteniamo una implicazione con antecedente falsa, che è sempre vera. Pertanto la formula data è sempre vera ed è chiusa; formule come questa si diranno **logicamente valide**.⁶

Definizione 2.23 (Formula chiusa). Una L -formula φ si dice **chiusa** se $\text{vl}(\varphi) = \emptyset$.

Definizione 2.24 (Formula logica valida). Una L -formula φ si dice **logicamente valida** se per ogni L -struttura $M = (D; i)$ e per ogni valutazione delle variabili $v : \text{Var} \rightarrow D$ vale $M \models \{v\}\varphi$.

⁶In generale assumeremo sempre che $D \neq \emptyset$ perché vogliamo che la formula $\forall x \varphi$ non sia sempre vera, e la formula $\exists x \varphi$ non sia sempre falsa.

Osservazione 2.25 (Indipendenza dalle variabili non libere) — Sia $M = (D; i)$ un modello e φ una L -formula, siano inoltre $v_1, v_2 : \text{Var} \rightarrow D$ tali che:

$$v_1|_{\text{vl}(\varphi)} = v_2|_{\text{vl}(\varphi)}$$

allora:

$$M \models \{v_1\}\varphi \iff M \models \{v_2\}\varphi$$

Dimostrazione. Procediamo per induzione strutturale.

L -termini Sia $t = x_k$, allora $\text{var}(t) = \{x_k\}$, segue che $\{v_1\}_M x_k = v_1(x_k) \stackrel{\text{hp.}}{=} v_2(x_k) = \{v_2\}_M x_k$. Sia ora $t = f(t_1, \dots, t_k)$, con $f \in F$ simbolo di funzione, $\text{ar}(f) = k$ e $t_i L$ -termini; dato che $\text{var}(t) = \bigcup_{i=1}^k \text{var}(t_i)$ segue dall'ipotesi che $v_1|_{\text{var}(t_i)} = v_2|_{\text{var}(t_i)}$ per ogni $i = 1, \dots, k$, a questo punto per ipotesi induttiva $\{v_1\}_M t_i = \{v_2\}_M t_i$, allora usando la definizione di interpretazione dei termini nella semantica di Tarski:

$$\begin{aligned} \{v_1\}_M f(t_1, \dots, t_k) &= i(f)(\{v_1\}_M t_1, \dots, \{v_1\}_M t_k) && (\text{hp. induttiva}) \\ &= i(f)(\{v_2\}_M t_1, \dots, \{v_2\}_M t_k) \\ &= \{v_2\}_M f(t_1, \dots, t_k) \end{aligned}$$

L -formule atomiche Sia $\varphi = \top$ o $\varphi = \perp$, allora $M \models \{v_1\}\varphi$ e $M \models \{v_2\}\varphi$ sono sempre soddisfatti o mai soddisfatti, quindi la tesi è banale. Se $\varphi = r(t_1, \dots, t_k)$, con $r \in R$ simbolo di relazione, $\text{ar}(r) = k$ e $t_i L$ -termini, allora $\text{vl}(\varphi) = \bigcup_{i=1}^k \text{var}(t_i)$, segue che $v_1|_{\text{var}(t_i)} = v_2|_{\text{var}(t_i)}$ per ogni $i = 1, \dots, k$, quindi per ipotesi induttiva $\{v_1\}_M t_i = \{v_2\}_M t_i$, allora usando la definizione di soddisfacibilità delle formule atomiche nella semantica di Tarski:

$$\begin{aligned} M \models \{v_1\}r(t_1, \dots, t_k) &\iff (\{v_1\}_M t_1, \dots, \{v_1\}_M t_k) \in i(r) && (\text{hp. induttiva}) \\ &\iff (\{v_2\}_M t_1, \dots, \{v_2\}_M t_k) \in i(r) \\ &\iff M \models \{v_2\}r(t_1, \dots, t_k) \end{aligned}$$

Infine se $\varphi = (t_1 = t_2)$, allora $\text{vl}(\varphi) = \text{var}(t_1) \cup \text{var}(t_2)$, segue che $v_1|_{\text{var}(t_i)} = v_2|_{\text{var}(t_i)}$, per cui per ipotesi induttiva $\{v_1\}_M t_i = \{v_2\}_M t_i$, ed usando ancora la definizione di soddisfacibilità della semantica di Tarski in questo caso si ottiene:

$$\begin{aligned} M \models \{v_1\}(t_1 = t_2) &\iff \{v_1\}_M t_1 = \{v_1\}_M t_2 && (\text{hp. induttiva}) \\ &\iff \{v_2\}_M t_1 = \{v_2\}_M t_2 \\ &\iff M \models \{v_2\}(t_1 = t_2) \end{aligned}$$

L -formule Sia ora $\varphi = \psi_1 \wedge \psi_2$ (o $\varphi = \psi_1 \vee \psi_2$ o $\varphi = \psi_1 \rightarrow \psi_2$), allora $\text{vl}(\varphi) = \text{vl}(\psi_1) \cup \text{vl}(\psi_2)$, segue che $v_1|_{\text{vl}(\psi_i)} = v_2|_{\text{vl}(\psi_i)}$, per cui per ipotesi induttiva $M \models \{v_1\}\psi_i \iff M \models \{v_2\}\psi_i$, allora per definizione di soddisfacibilità nella semantica di Tarski:

$$\begin{aligned} M \models \{v_1\}(\psi_1 \wedge \psi_2) &\iff M \models \{v_1\}\psi_1 \wedge M \models \{v_1\}\psi_2 && (\text{hp. induttiva}) \\ &\iff M \models \{v_2\}\psi_1 \wedge M \models \{v_2\}\psi_2 \\ &\iff M \models \{v_2\}(\psi_1 \wedge \psi_2) \end{aligned}$$

e analogamente per \vee e \rightarrow . Sia ora $\varphi = \forall x_k \psi$, allora $\text{vl}(\varphi) = \text{vl}(\psi) \setminus \{x_k\}$, assumiamo che $M \models \{v_1\} \forall x_k \psi$, che equivale per definizione a $\forall a \in D M \models \{v_1[a/x_k]\}\psi$

e dimostriamo che $M \models \{v_2\} \forall x_k \psi$. Fissiamo $a \in D$, allora per ipotesi induttiva si ha che $M \models \{v_1[a/x_k]\} \psi \iff M \models \{v_2[a/x_k]\} \psi$, infatti $v_1[a/x_k]|_{\text{vl}(\psi)} = v_2[a/x_k]|_{\text{vl}(\psi)}$ (ovvio in x_k perché vengono entrambi a , e per tutte le altre variabili vale l'ipotesi) e l'uguaglianza segue dall'ipotesi induttiva; a questo punto abbiamo che $\forall a \in D M \models \{v_1[a/x_k]\} \psi \iff M \models \{v_1[a/x_k]\} \psi$, e per definizione di semantica di Tarskiabbiamo: $M \models \{v_1\} \forall x_k \psi \iff M \models \{v_2\} \forall x_k \psi$.

Analogamente se $\varphi = \exists x_k \psi$, allora $\text{vl}(\varphi) = \text{vl}(\psi) \setminus \{x_k\}$, assumiamo che $M \models \{v_1\} \exists x_k \psi$, che equivale per definizione a $\exists a \in D M \models \{v_1[a/x_k]\} \psi$, fissato un $a \in D$ per cui M soddisfa $\{v_1[a/x_k]\} \psi$, si ha che per ipotesi induttiva $M \models \{v_1[a/x_k]\} \psi \iff M \models \{v_2[a/x_k]\} \psi$, infatti $v_1[a/x_k]|_{\text{vl}(\psi)} = v_2[a/x_k]|_{\text{vl}(\psi)}$ (come prima), a questo punto, ancora come prima abbiamo che $\exists a \in D M \models \{v_1[a/x_k]\} \psi \iff M \models \{v_1[a/x_k]\} \psi$, cioè $M \models \{v_1\} \exists x_k \psi \iff M \models \{v_2\} \exists x_k \psi$.

□

Corollario 2.26 (Soddisfacibilità delle formule chiuse)

Se φ è una formula **chiusa**, allora φ vale in qualche contesto (interpretazione) se e solo se vale in ogni contesto. Ossia, data una qualunque valutazione delle variabili v :

$$M \models \{v\} \varphi \iff \forall v : \text{Var} \rightarrow D M \models \{v\} \varphi$$

Segue che per verificare che una formula chiusa sia logicamente valida è sufficiente trovare, per ogni modello, un'interpretazione in cui sia valida.

Notazione 2.27 — Scriviamo che $M \models \varphi$, senza specificare il contesto, per dire che M soddisfa φ in ogni contesto, i.e.:

$$M \models \varphi \stackrel{\text{def}}{\iff} \forall v : \text{Var} \rightarrow D M \models \{v\} \varphi$$

Osservazione 2.28 (Soddisfacibilità per ogni interpretazione) — La scrittura $M \models \varphi$ ha senso anche se φ non è una formula chiusa. In questo caso, se $\text{vl}(\varphi) = \{\alpha_1, \dots, \alpha_n\}$ vale che:

$$M \models \varphi \stackrel{\text{def}}{\iff} M \models \forall \alpha_1, \dots, \forall \alpha_n \varphi$$

e quest'ultima è una formula chiusa. Infatti, più in generale si ha che:

$$M \models \psi \stackrel{\text{def}}{\iff} M \models \forall x_k \psi$$

Esercizio 2.29. Verificare l'osservazione precedente.

§2.3 Sostituzioni

Questa breve sezione esiste per accomodare una scomodità legata alla nostra ostinazione di usare, come formule, delle liste di simboli. Finché ci limitiamo, per esempio, alle identità algebriche, è chiaro che possiamo sostituire un termine qualunque, al posto di una variabile qualunque, in un'identità [logicamente] valida, ottenendo ancora un'identità valida. Per esempio da $(x+y)(x-y) = x^2 - y^2$, scrivendo al posto di x , il termine $1+y$ ottengo $(1+y+y)(1+y-y) = (1+y)^2 - y^2$, che è ancora un'identità valida. Il fatto che y compaia sia nella identità di partenza sia nel termine sostitutivo non compromette la validità di questo procedimento. Se tento lo stesso procedimento ad esempio con

la formula $\exists y \ x < y$, valida nella struttura $(\mathbb{Q}, 1, +, <)$, e sostituisco $1 + y$ al posto di x , ottengo $\exists y \ 1 + y < y$, che non è più logicamente valida (non vale più in qualsiasi L -struttura). La radice del guaio è fin troppo ovvia: la formula $\exists y \ x < y$ dice che c'è un y , che può dipendere da x , che si trova rispetto a x in una certa situazione. Scrivendo $1 + y$ al posto di x , impongo anche una dipendenza di x da y , creando così un ciclo di dipendenze. È vero che dato un x posso trovare un y , ma non necessariamente questo y soddisfa il vincolo ulteriore di chiudere i cicli. Se scrivessi la formula così (cromaticamente):

$$\exists y \ 1 + y < y$$

oppure così (biscromaticamente): non ci sarebbero problemi, perché il y (o simbolo) che

$$\exists \bullet \overline{1+y < \bullet}$$

compare nella formula non è lo stesso y che compare nel termine $1 + y$. Volendo tuttavia utilizzare gli stessi simboli per le variabili legate per le variabili libere - convenzione che ha i suoi vantaggi - si cade occasionalmente, ma inevitabilmente nel [problema delle catture delle variabili](#).

Come ne usciamo? Intanto rallegriamoci! Per gli informatici è peggio: il λ -calcolo **vive** di sostituzioni, ed è lì che il male ha messo radici. Noi, ce la caveremo semplicemente vietando le sostituzioni insalubri, cosa che, nel nostro contesto, non ha controindicazioni.

Definizione 2.30 (Sostituibilità). Sia φ una L -formula e sia t un L -termine. Diciamo che t è **sostituibile** per x_k in φ se nessuna occorrenza libera di x_k in φ si trova in una sottoformula del tipo $\forall \alpha \psi$ o $\exists \alpha \psi$ con $\alpha \in \text{var}(t)$. Più formalmente, usando la ricorsione strutturale diciamo che t è **sostituibile** per x_k in φ se:

- φ è atomica;
- $\varphi = \neg \psi$ e t è sostituibile per x_k in ψ ;
- $\varphi = \psi_1 \wedge \psi_2$ (o $\varphi = \psi_1 \vee \psi_2$ o $\varphi = \psi_1 \rightarrow \psi_2$) e t è sostituibile per x_k in ψ_1 e in ψ_2 ;
- $\varphi = \forall x_i \psi$ (o $\varphi = \exists x_i \psi$) e si verifica uno dei casi seguenti: o $x_k \in \text{vl}(\varphi)$, $x_i \notin \text{var}(t)$ (cioè la variabile quantifica in φ non appare tra le variabili libere di t , se fosse diversamente, tale variabile verrebbe quantificata - catturata - a sua volta) e t è sostituibile per x_k in ψ ; oppure $x_k \notin \text{vl}(\varphi)$.

Capiamo prima la [definizione informale](#). Una occorrenza di un simbolo α in una stringa s è un indice i tale che $s_i = \alpha$. Fra le occorrenze del simbolo x_k in φ ce ne sono alcune **legate**, quelle che fanno parte di una sottoformula del tipo $\forall x_k \dots$ o $\exists x_k \dots$, e le altre sono **libere**. Pedantemente, i è un'occorrenza legata se ci sono j_1 e j_2 con $j_1 \leq i \leq j_2$ tali che la sottostringa di φ costituita dai caratteri che vanno dal j_1 -esimo al j_2 -esimo è una sottoformula che inizia per $\forall x_k$ o $\exists x_k$.

Per esempio:

$$\forall y \ \forall z (y \cdot z = x \rightarrow \exists x \ \exists t \ t = x + x + x \wedge y \cdot s(t) = y \cdot y + t)$$

ha una occorrenza libera di x (la prima), mentre le altre sono legate. È chiaro che le variabili libere di una formula sono quelle che hanno almeno una occorrenza libera (in questo caso c'è solo la prima x). In questa formula $s(y)$ e $z \cdot z$ NON sono sostituibili per x , mentre $t + t$ lo è.

Esercizio 2.31. Nella struttura $(\mathbb{N}, s, +, \cdot)$, dove s denota il successore, cosa significa quel delirio sopra?

Esercizio 2.32 (Difficile). Riesci a rimpiazzare 3 con 10 nella formula sopra?⁷

Esercizio 2.33. Convinciti della definizione formale.

Osservazione 2.34 (Le variabili non libere sono sempre sostituibili) — Se $x_k \notin \text{vl}(\varphi)$, allora qualunque L -termine t è sostituibile per x_k .

Osservazione 2.35 (Le costanti possono essere sempre sostituite) — Se c è un simbolo di funzione di arietà 0 (costante), allora è sostituibile per x_k in qualunque L -formula φ .^a

^aLa ragione è che $\text{var}(c) = \emptyset$, quindi non può capitare che $x_i \in \text{var}(c)$.

Osservazione 2.36 (Gli L -termini semplici possono essere sempre sostituiti) — $f(x_k)$ è sostituibile per x_k in qualunque L -formula φ .^a

^aInfatti $\text{var}(f(x_k)) = \{x_k\}$, quindi non può capitare che $x_i \in \text{var}(f(x_k))$ (e se $x_i = x_k$ allora x_k non sarebbe un'occorrenza libera, per cui saremmo nel caso $x_k \notin \text{vl}(\varphi)$).

Bene, sappiamo cosa significa che un termine è sostituibile, ma come si fanno le sostituzioni?

Definizione 2.37 (Sostituzione di una variabile libera con un L -termine). Sia φ una L -formula e t un L -termine **sostituibile** per x_k in φ . Denotiamo con $\varphi[t/x_k]$ la formula ottenuta rimpiazzando tutte le occorrenze libere di x_k in φ con t . Più formalmente, per ricorsione strutturale:

L -termini Se $t = x_i$, allora:

$$x_i[t/x_k] = \begin{cases} t & \text{se } i = k \\ x_i & \text{se } i \neq k \end{cases}$$

Se $t = f(t_1, \dots, t_n)$, con $f \in F$ simbolo di funzione, allora:

$$f(t_1, \dots, t_n)[t/x_k] = f(t_1[t/x_k], \dots, t_n[t/x_k])$$

L -formule atomiche Se $\varphi = \top$ o $\varphi = \perp$, allora $\varphi[t/x_k] = \varphi$. Se $\varphi = r(t_1, \dots, t_n)$, con $r \in R$ simbolo di relazione, allora:

$$r(t_1, \dots, t_n)[t/x_k] = r(t_1[t/x_k], \dots, t_n[t/x_k])$$

Se $\varphi = (t_1 = t_2)$, allora $(t_1 = t_2)[t/x_k] = (t_1[t/x_k] = t_2[t/x_k])$.

L -formule Se $\varphi = \neg\psi$, allora $\varphi[t/x_k] = \neg(\psi[t/x_k])$. Se $\varphi = \psi_1 \wedge \psi_2$ (o $\varphi = \psi_1 \vee \psi_2$ o $\varphi = \psi_1 \rightarrow \psi_2$), allora:

$$(\psi_1 \wedge \psi_2)[t/x_k] = \psi_1[t/x_k] \wedge \psi_2[t/x_k]$$

e similmente negli altri casi. Se $\varphi = \forall x_i \psi$ (o $\varphi = \exists x_i \psi$), allora:

$$(\forall x_i \psi)[t/x_k] = \begin{cases} \forall x_i (\psi[t/x_k]) & \text{se } x_k \neq x_i \\ \forall x_i \psi & \text{se } x_k = x_i \end{cases}$$

e similmente nel caso esistenziale.

⁷A fine corso sarà facile, ma per ora è difficile.

Nota 2.38 — Quando scriviamo $\varphi[t/x_k]$ assumiamo che t è sostituibile per x_k in φ . La scrittura non ha senso altrimenti. Per esempio, detta φ la formula di prima:

$$\forall y \forall z (y \cdot z = \underbrace{x}_{\text{libera}} \rightarrow \exists x \exists t t = \underbrace{x+x+x}_{\text{legate}} \wedge y \cdot s(t) = y \cdot y + t)$$

La formula $\varphi[t + t/x]$ è:

$$\forall y \forall z (y \cdot z = t + t \rightarrow \exists x \exists t t = x + x + x \wedge y \cdot s(t) = y \cdot y + t)$$

che ragionevolmente equivale a:

$$\forall y \forall z (y \cdot z = t + t \rightarrow \exists x \exists n n = x + x + x \wedge y \cdot s(n) = y \cdot y + n)$$

tuttavia quest'ultima non si ottiene come sostituzione secondo la definizione precedente in quanto n non è sostituibile per t in φ .

Esercizio 2.39 (Sostituzione e valutazione delle variabili commutano). Vale il seguente fatto: $M \models \{v\}\varphi[t/x_k] \iff M \models \{v[t/x_k]\}\varphi$.

Ossia sostituire t al posto di x_k ha il medesimo effetto che valutare t e assegnare, nell'ambiente, il valore di t alla variabile x_k .

Questo asserto si dimostra precisamente come l'osservazione che $M \models \{v_1\}\varphi \iff M \models \{v_2\}\varphi$ se v_1 e v_2 coincidono sulle variabili libere di φ , ma con più pasticcio di notazioni. Ci servirà per giustificare una delle regole di deduzione.

§2.4 Teorie

Definizione 2.40 (*L-teoria*). Una ***L-teoria*** è un insieme di *L*-formule.

Definizione 2.41 (Modello). Una *L*-struttura M si dice **modello** di una *L*-teoria T se $\forall \varphi \in T$ si ha $M \models \varphi$.⁸

Definizione 2.42 (Conseguenza logica). La *L*-formula φ è **conseguenza logica** della *L*-teoria T , e si scrive $T \models \varphi$, se per ogni modello M di T vale $M \models \varphi$.

Nota 2.43 (Conseguenza logica del vuoto) — Il simbolo \models si può usare anche con la teoria vuota a sinistra. Si scrive $\models \varphi$, e ciò equivale a dire che φ è logicamente valida.

Definizione 2.44 (Coerenza). Una *L*-teoria T è **coerente** se ha un modello.

Osservazione 2.45 (Caratterizzazione della coerenza) — T è coerente se e solo se $T \not\models \perp$.

Dimostrazione. Vediamo le due implicazioni.

⇒ Se T è coerente, allora ha un modello M , per cui deve valere che $M \models \perp$ per definizione di conseguenza logica, ma questo non può essere, per definizione di soddisfacibilità, in quanto \perp non è mai soddisfatta (è vero che $\neg M \models \perp$).

⁸Come già osservato con la seconda cosa si intende che, detto $\text{vl}(\varphi) = \{\alpha_1, \dots, \alpha_n\}$, si ha $M \models \forall \alpha_1 \dots \forall \alpha_n \varphi$.

\Leftarrow Se $T \not\models \perp$, allora esiste un modello M di T tale che $M \not\models \perp$, ma questo è sempre vero per definizione di soddisfabilità nella semantica di Tarski, per cui M è un modello di T e quindi T è coerente.

□

Definizione 2.46 (Completezza). La L -teoria T è **completa** se, per ogni L -formula **chiusa** φ , vale una e una sola delle seguenti: $T \models \varphi$ oppure $T \models \neg\varphi$.

Osservazione 2.47 (Completezza \implies coerenza) — Se una L -teoria T è completa, allora è coerente.

Dimostrazione. Per definizione $T \models \top$, infatti ogni modello di T soddisfa \top per definizione di soddisfabilità nella semantica di Tarski, dunque, per completezza, vale necessariamente che $T \not\models \neg\top = \perp$, per cui T è coerente per la caratterizzazione vista prima. □

§2.5 Esempi di teorie

È facile costruire una teoria incoerente, per esempio prendendo come **assioma** il falso: $T = \{\perp\}$. Non sempre, però, è facile distinguere l'incoerenza.

Esercizio 2.48 (Un esempio di teoria incoerente). Dimostra che la teoria, nel linguaggio $\{f, \alpha, \beta\}$ dove f è un simbolo di funzione binaria e α, β sono simboli di costante, data dagli assiomi che seguono:

$$\begin{aligned} & \forall x \forall y \exists z \forall t f(x, f(y, t)) = f(z, t) \\ & \forall x f(\alpha, x) = f(x, x) \\ & \forall x \neg(f(\beta, x) = x) \end{aligned}$$

non è coerente.

Per un esempio coerente possiamo considerare la teoria dei gruppi, i cui modelli saranno tutti e soli i gruppi.

Esempio 2.49 (Teoria dei gruppi)

Consideriamo il linguaggio $L_{\text{gruppi}} = \{e, \dots \cdot \dots, \dots^{-1}\}$, e:

$$\begin{aligned} T_{\text{gruppi}} = & \{ \forall x \forall y \forall z (x \cdot (y \cdot z) = (x \cdot y) \cdot z), \\ & \forall x e \cdot x = x \quad \forall x x \cdot e = x, \\ & \forall x x \cdot x^{-1} = e \quad \forall x x^{-1} \cdot x = e \} \end{aligned}$$

Esercizio 2.50. Convinciti del fatto che una struttura \mathcal{G} è un modello di T_{gruppi} se e solo se \mathcal{G} è un gruppo.

La teoria dei gruppi è chiaramente coerente: basta esibire un gruppo qualunque per avere un modello. Questa non è una teoria completa, infatti, ci sono gruppi in cui vale $\varphi = \forall x \forall y x \cdot y = y \cdot x$ e ce ne sono in cui vale $\neg\varphi$ (quindi viene meno la definizione di completezza di una L -teoria).

Un esempio di teoria completa si può ottenere considerando tutte le formule vere in una cera struttura.

Definizione 2.51 (Teoria completa di una struttura). Data una L -struttura M , definiamo la **teoria completa di M** , denotata $\text{Th}(M)$ è l'insieme di tutte le L -formule φ tali che $M \models \varphi$.

Osservazione 2.52 — $\text{Th}(M)$ è una L -teoria completa.

Si potrebbe pensare a prima vista che una teoria completa caratterizzi un certo modello, a meno di isomorfismi. Non è così, se non nel caso finito.

Definizione 2.53 (Morfismi di strutture). Date due L -strutture $M = (D; i)$ e $M' = (D'; i')$ un **morfismo di strutture** $F : M \rightarrow M'$ è una funzione $F : D \rightarrow D'$ tale che:

(i) Per ogni simbolo di relazione r , e $(x_1, \dots, x_{\text{ar}(r)}) \in D^{\text{ar}(r)}$ si ha che:

$$(x_1, \dots, x_{\text{ar}(r)}) \in r_M \implies (F(x_1), \dots, F(x_{\text{ar}(r)})) \in r_{M'}$$

(ii) Per ogni simbolo di funzione f , e $(x_1, \dots, x_{\text{ar}(f)}) \in D^{\text{ar}(f)}$ si ha che:

$$F \circ f_M(x_1, \dots, x_{\text{ar}(f)}) = f_{M'}(F(x_1), \dots, F(x_{\text{ar}(f)}))$$

Definizione 2.54 (Immersioni e isomorfismi di strutture). Un morfismo di L -strutture $F : M \rightarrow M'$ si dice **immersione** se F , come funzione tra i domini, è iniettiva.

Un'immersione F è un **isomorfismo** se F è altresì surgettiva, o F^{-1} è, a sua volta, un morfismo di strutture.

Esercizio 2.55 (Teorie complete e modelli finiti). Dimostra che, se T è completa e M è un modello di T , avente dominio finito, allora tutti i modelli di T sono isomorfi a M .

Se T ha un modello infinito, però, ne ha almeno uno per ogni cardinalità maggiore o uguale a $|L|$. Questo risultato, che vedremo più avanti, preclude una volta per tutte la possibilità di caratterizzare una struttura per mezzo di una teoria del primo ordine (nel senso che due strutture con la stessa teoria non è detto che siano isomorfe).

Quanto a caratterizzare, il meglio che possiamo sperare è di esibire teorie complete, ossia caratterizzare non una struttura, bensì l'insieme degli enunciati veri in una struttura. La teoria degli ordini totali, densi e senza estremi è un esempio di teoria completa descritta esplicitamente che ha modelli infiniti.

Esempio 2.56 (Teoria degli ordini totali, densi e senza estremi)

Consideriamo il linguaggio $L_{\text{otdse}} = \{\langle\}$, e la teoria:

$$\begin{aligned} T_{\text{otdse}} = & \{ \forall x \forall y \forall z x < y \wedge y < z \rightarrow x < z, & (\text{transitività}) \\ & \forall x \neg x < x, & (\text{irriflessività}) \\ & \forall x \forall y x < y \vee x = y \vee y < x, & (\text{totalità}) \\ & \forall x \forall y x < y \rightarrow \exists z x < z \wedge z < y, & (\text{densità}) \\ & \forall x \exists y x < y, & (\text{assenza di massimo}) \\ & \forall x \exists y y < x & (\text{assenza di minimo}) \} \end{aligned}$$

Esercizio 2.57 (Completezza di T_{otdse}). Leggi nella prossima sezione come dimostrare che T_{otdse} è completa.

§2.6 PA e Q di Robinson

Altre due teorie rilevanti per questo corso sono due sottoinsiemi di $\text{Th}(\mathbb{N}; 0, s, +, \cdot)$, dove s rappresenta la funzione successore: l'**aritmetica di Peano (PA)** e la **teoria Q di Robinson**. Entrambe sono teorie nel **linguaggio dell'aritmetica** $L_{\text{arit}} = \{0, s, +, \cdot\}$.

Definizione 2.58 (PA e Q di Robinson). Le L_{arit} -teorie PA e Q hanno in comune i seguenti assiomi:

- Q1** $\forall x \neg s(x) = 0$ (0 non successore);
- Q2** $\forall x \forall y s(x) = s(y) \rightarrow x = y$ (iniettività del successore);
- Q3** $\forall x x + 0 = x$ (def. ricorsiva somma);
- Q4** $\forall x \forall y x + s(y) = s(x + y)$ (def. ricorsiva somma);
- Q5** $\forall x x \cdot 0 = 0$ (def. ricorsiva prodotto);
- Q6** $\forall x \forall y x \cdot s(y) = (x \cdot y) + x$ (def. ricorsiva prodotto).

A Q1 – 6, la teoria Q aggiunge il seguente assioma:

- Q7** $\forall x x = 0 \vee \exists y s(y) = x$ (ogni numero eccetto 0 è successore).

Mentre PA aggiunge a Q1 – 6 il seguente **schema di induzione**:

$$\mathbf{I}_\varphi (\varphi[0/x_k] \wedge \forall x_k (\varphi \rightarrow \varphi[s(x_k)/x_k])) \rightarrow \forall x_k \varphi.$$

Ossia PA contiene una formula I_φ per ogni possibile L_{arit} -formula φ e ogni possibile variabile x_k .

Notazione 2.59 — Se indichiamo una formula qualunque con la scrittura $\varphi(x)$, è per dire che, quando poi scriviamo $\varphi(t)$, intenderemo $\varphi[t/x]$. Così lo schema di induzione si può scrivere più familiarmente:

$$(\varphi(0) \wedge \forall x \varphi(x) \rightarrow \varphi(s(x))) \rightarrow \forall x \varphi(x)$$

al variare di $\varphi(x)$ fra tutte le formule e di x fra tutte le variabili.

Esercizio 2.60 (\mathbb{N} modella PA e Q). Convinciti del fatto che $\mathbb{N} \models \text{PA}$ e quindi anche $\mathbb{N} \models \text{Q}$.

Esercizio 2.61 (Q7 è conseguenza logica di PA). Dimostra che $\text{PA} \models \text{Q7}$, quindi tutti i modelli di PA sono modelli di Q.

Esercizio 2.62 (Q non è completa). Trova un modello di Q che non è un modello di PA e deducine che Q non è completa.

Per l'esercizio precedente, Q non è completa, ma, a prima vista, si potrebbe pensare PA lo sia. Infatti è ben noto che il principio di induzione:

$$\forall X \subseteq \mathbb{N} (0 \in X \wedge \forall n \in \mathbb{N} n \in X \rightarrow s(n) \in X) \rightarrow X = \mathbb{N}$$

caratterizza \mathbb{N} a meno di isomorfismi. Tuttavia non è così, e PA NON è completa: questo è il famoso **primo teorema di incompletezza di Gödel**. Com'è possibile?

Il guaio sta nel fatto che il principio di induzione scritto qua sopra, che potremmo chiamare **induzione al secondo ordine**, fa riferimento ad ogni possibile sottoinsieme X di \mathbb{N} . Lo schema di induzione di PA, d'altro canto, lavora solo sui sottoinsiemi $X = \{n | \varphi(n)\}$ per qualche L_{arit} -formula φ , e questi sono molti meno di tutti i sottoinsiemi di \mathbb{N} , perché c'è solo una quantità numerabile di formule. La dimostrazione dell'incompletezza di PA non è banale, e si vedrà in questo corso.

§3 Eliminazione dei quantificatori, forme normali ed ultrafiltri

§3.1 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 $T \models (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$, 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 (T_{oldse} ha l'eliminazione dei quantificatori)

La teoria degli ordini lineari densi senza estremi T_{oldse} ha la proprietà dell'eliminazione dei quantificatori.

Corollario 3.4 (Completezza di T_{oldse})

La teoria T_{oldse} è completa.

Vediamo la dimostrazione del corollario.

Dimostrazione. T_{oldse} è coerente perché $\mathbb{Q} \models T_{\text{oldse}}$. Inoltre, nel linguaggio di T_{oldse} , non ci sono simboli di costante (c'è solo un simbolo di relazione), quindi le uniche formule chiuse senza quantificatori sono combinazioni booleane di \top e \perp . È immediato che queste formule hanno un valore di verità definito. Pertanto come osservato sopra, T_{oldse} è completa rispetto alle formule chiuse e senza quantificatori, e per la proposizione precedente questo è sufficiente a dire che T_{oldse} è completa. \square

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

Nota 3.5 (L'eliminazione dei quantificatori dipende dal linguaggio) — La scelta del linguaggio è cruciale per l'eliminazione dei quantificatori. Per esempio la teoria $\text{Th}(\mathbb{R}; 0, 1, +, \cdot)$ può esprimere le medesime proprietà di $\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$, perché $x < y$ equivale a $x \neq y \wedge \exists z x + z \cdot z = y$. Tuttavia, senza il simbolo $<$, non c'è modo di eliminare il quantificatore esistenziale in $\exists y x = y \cdot y$ (esercizio: perché?). In un linguaggio opportuno, ogni teoria ha l'eliminazione dei quantificatori: basta infatti aggiungere un simbolo di relazione \mathcal{R}_φ per ogni formula $\varphi(x_1, \dots, x_k)$ con l'assioma $\varphi(x_1, \dots, x_k) \leftrightarrow \mathcal{R}_\varphi(x_1, \dots, x_k)$.

Questa nuova teoria ha ovviamente l'eliminazione dei quantificatori, che però è inutile, perché decidere le formule atomiche nel linguaggio espanso è tanto complesso quanto decidere le formule della teoria di partenza.

Vediamo la soluzione dell'esercizio nell'osservazione.

Soluzione. Sia $\varphi := \exists y x = y \cdot y$ e supponiamo che esista ψ , senza quantificatori, $\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$ -equivalente a φ nel linguaggio $L = \{0, 1, +, \cdot\}$. Una tale ψ , essen-

do senza quantificatori e con una sola variabile libera, può essere scritta in DNF come una disgiunzione di congiunzioni di formule del tipo $p(x) = 0$ o $p(x) \neq 0$, dove $p(x) \in \mathbb{R}[x]$. Ora \mathbb{R} è chiaramente un modello di $\text{Th}(\mathbb{R}; 0, 1, +, \cdot)$, per cui se ψ e φ sono equivalenti, allora devono essere soddisfatte dagli stessi valori in \mathbb{R} , in particolare φ è soddisfatta da tutti i reali in $[0, +\infty)$, mentre ψ corrisponde ad un'unione finita di intersezioni di chiusi di Zariski e loro complementari, ora se in quest'unione c'è almeno un cofinito, l'unione è cofinita e quindi non può essere uguale a $[0, +\infty)$, altrimenti l'unione è finita e quindi non può essere uguale a $[0, +\infty)$. \square

Osservazione 3.6 (Sostituzione di formule equivalenti) — 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 $T \models \psi_1 \leftrightarrow \psi_2$.

Dimostrazione. Procediamo per induzione strutturale su ψ_1 .

ψ_1 atomica In tal caso necessariamente $\psi_1 = \varphi_1$, per cui la tesi segue subito dall'ipotesi.

$\psi_1 = \neg\varphi_1$ Allora, per ipotesi si ha che $T \models \varphi_1 \leftrightarrow \varphi_2$, inoltre è sempre vero che $\models (\varphi_1 \leftrightarrow \varphi_2) \rightarrow (\neg\varphi_1 \leftrightarrow \neg\varphi_2)$, quindi $T \models \neg\varphi_1 \leftrightarrow \neg\varphi_2$ (volendo per transitività ogni volta che fisso un modello ed uso la semantica di Tarski).

$\psi_1 = \varphi_1 * \theta$ Con $* \in \{\wedge, \vee, \rightarrow\}$, per ipotesi induttiva si ha che $T \models \varphi_1 \leftrightarrow \varphi_2$ dove φ_2 è la formula ottenuta da φ_1 sostituendo una sottoformula (WLOG ho assunto di farlo in φ_1), a questo punto dalla semantica di Tarski si deduce che:

$$\begin{aligned} T \models \varphi_1 * \theta &\iff T \models \varphi_1 * T \models \theta \\ &\iff T \models \varphi_2 * T \models \theta \iff T \models \varphi_2 * \theta \end{aligned}$$

ovvero $T \models \varphi_1 * \theta \iff T \models \varphi_2 * \theta$, e questo per la semantica di Tarski (la sto usando ma soprassedendo sempre sul modello specifico fissato) equivale a dire che $T \models (\varphi_1 * \theta) \leftrightarrow (\varphi_2 * \theta)$.

$\psi_1 = \forall x_k \varphi_1$ Per ipotesi induttiva si ha che $T \models \varphi_1 \leftrightarrow \varphi_2$ dove φ_2 è la formula ottenuta da φ_1 sostituendo una sottoformula equivalente per T . Ora fissato un modello $M = (D; i)$ di T e una valutazione v delle variabili, attraverso la semantica di Tarski si ha che:

$$\begin{aligned} M \models \{v\} \forall x_k \varphi_1 &\iff \forall a \in D M \models \{v[a/x_k]\} \varphi_1 \\ &\iff \forall a \in D M \models \{v[a/x_k]\} \varphi_2 \\ &\iff M \models \{v\} \forall x_k \varphi_2 \end{aligned} \tag{*}$$

dove in (*) abbiamo semplicemente usato che l'equivalenza vale per ogni modello e valutazione delle variabili. L'equivalenza ottenuta corrisponde, per la semantica di Tarski, a $T \models \forall x_k \varphi_1 \leftrightarrow \forall x_k \varphi_2$. In maniera analoga si procede con il quantificatore esistenziale.

\square

§3.2 CNF e DNF

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

Definizione 3.7 (CNF e DNF). Una formula senza quantificatori φ è in **forma normale congiuntiva (CNF)** se:

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

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:

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

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

Lemma 3.8 (Ogni formula senza quantificatori è equivalente ad una in CNF e DNF)

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).

Dimostrazione. Ci sono almeno due vie possibili:

1. Considerare ogni possibile assegnazione dei valori di verità alle formule atomiche che compaiono in φ . Le asserzioni che rendono φ vera danno i disgiunti (**implicanti primi**) della DNF, quelle che rendono φ falsa danno i congiunti (**clausole**) della CNF.
2. Per induzione strutturale dimostra che se φ_1 e φ_2 possono essere espresse in CNF e DNF, allora anche $\neg\varphi_1$, $\varphi_1 \wedge \varphi_2$ possono essere espresse in CNF e DNF. Osservare inoltre che i connettivi \neg e \wedge possono essere usati per esprimere tutti gli altri connettivi.

□

Dimostriamo ora finalmente che T_{oldse} ha l'eliminazione dei quantificatori.

Dimostrazione. Procediamo per induzione strutturale, sia φ libera dai quantificatori e verifichiamo che $\exists x_k \varphi$ è T_{oldse} -equivalente ad una formula senza quantificatori. Non sarà necessario fare lo stesso anche con il quantificatore universale in quanto $\models \forall x_k \varphi \leftrightarrow \neg(\exists x_k \neg\varphi)$ per cui ci basta poter eliminare \exists . Per ipotesi induttiva, usando il lemma sopra, posso scrivere φ in DNF:

$$\varphi = \varphi_1 \vee \dots \vee \varphi_m$$

dove ogni φ_i è un implicante primo (congiunzione di formule atomiche o negazioni di esse). Si osserva che $\models \exists x_k \varphi \leftrightarrow (\exists x_k \varphi_1) \vee \dots \vee (\exists x_k \varphi_m)$, per cui basta eliminare \exists da $\exists x_k \varphi_i$. Posso assumere che in φ_i non ci siano formule chiuse, perché posso portarle fuori dall'esistenziale, ora le formule non chiuse atomiche sono solo del tipo $x_j < x_k$ o $\neg(x_j < x_k)$ che posso abbreviare con $x_j \geq x_k$, in quanto $T_{\text{oldse}} \models \neg(x_j < x_k) \leftrightarrow x_j \geq x_k$.

Infine scrivo $\varphi_i = \varphi_1 \wedge \varphi_2$, dove metto in φ_1 tutte le formule in cui x_k compare a destra, e in φ_2 tutte le formule in cui x_k compare a sinistra:

$$\begin{aligned}\varphi_1 &= v_1 \underbrace{\bullet}_{\alpha_{11}} \underbrace{x_k}_{\leq} \wedge \dots \wedge v_m \underbrace{\bullet}_{\alpha_{1m}} \underbrace{x_k}_{\leq} && \text{con } v_i \text{ variabili diverse da } x_k \\ \varphi_2 &= x_k \underbrace{\bullet}_{\alpha_{21}} \underbrace{w_1}_{\leq} \wedge \dots \wedge x_k \underbrace{\bullet}_{\alpha_{2n}} \underbrace{w_n}_{\leq} && \text{con } w_i \text{ variabili diverse da } x_k\end{aligned}$$

A questo punto definisco $\psi := \bigwedge_{i=1 \dots m} \bigwedge_{j=1 \dots n} \psi_{i,j}$, dove:

$$\psi_{i,j} := \begin{cases} v_i \leq w_j & \text{se } \alpha_{1i} = v_i \leq x_k \text{ e } \alpha_{2j} = x_k \leq w_j \\ v_i < w_j & \text{se almeno uno dei segni di } \alpha_{1i} \text{ e } \alpha_{2j} \text{ è } < \end{cases}$$

Se uno tra m o n è 0, poniamo $\psi := \top$. Claimmiamo ora che $T_{\text{oldse}} \models (\exists x_k \varphi_i) \leftrightarrow \psi$. Vediamo le due implicazioni.

- Sia M un modello di T_{oldse} e sia v una valutazione tali che $M \models \{v\} \exists x_k \varphi_i$, allora per gli assiomi di T_{oldse} , ovvero la transitività, segue subito che tutte le $\psi_{i,j}$ valgono in M , e quindi per la semantica di Tarski anche ψ vale in M .
- ← Sia M un modello di T_{oldse} e v una valutazione delle variabili in cui vale ψ . Se $m \neq 0$ esiste una $\alpha_{1\bar{i}}$ più restrittiva delle altre (ovvero quella con $v(v_i)$ più grande, e posso sceglierlo nel mio modello), chiamo $\bar{\varphi}_1 := v_{\bar{i}} \bullet \underbrace{x_k}_{\leq}$ questa condizione, se $m = 0$ pongo $\bar{\varphi}_1 := \top$. Similmente, se $n \neq 0$ esiste una $\alpha_{2\bar{j}}$ più restrittiva delle altre (ovvero quella con w_j più piccola, e posso sceglierlo nel mio modello perché soddisfa gli assiomi di ordine totale), chiamo $\bar{\varphi}_2 := x_k \bullet \underbrace{w_{\bar{j}}}_{\leq}$ questa condizione, se $n = 0$ pongo $\bar{\varphi}_2 := \top$. Segue immediatamente che:

$$M \models \{v\} \forall x_k \bar{\varphi}_1 \rightarrow \varphi_1 \quad \text{e} \quad M \models \{v\} \forall x_k \bar{\varphi}_2 \rightarrow \varphi_2$$

di conseguenza $M \models \{v\} \forall x_k (\bar{\varphi}_1 \wedge \bar{\varphi}_2) \rightarrow \varphi_i$, quindi per verificare che $M \models \{v\} \exists x_k \varphi_i$, è sufficiente vedere che:

$$M \models \{v\} \psi \rightarrow \exists x_k \bar{\varphi}_1 \wedge \bar{\varphi}_2$$

Distinguiamo due casi:

- Se almeno uno tra m o n è 0, allora rimane solo una delle due condizioni, quindi solo una disegualanza da soddisfare almeno una volta, e questo è vero in M perché è un ordine senza estremi.
- Se sia m che n sono diversi da 0, allora, per ipotesi so che $M \models \{v\} \psi_{\bar{i},\bar{j}}$, con $\psi_{\bar{i},\bar{j}} = v_{\bar{i}} \bullet \underbrace{w_{\bar{j}}}_{\leq}$, restano due sottocasi.
 - Se sia $\bar{\varphi}_1$ che $\bar{\varphi}_2$ usano \leq , allora posso scegliere come x_k uno tra $v(v_{\bar{i}})$ e $v(w_{\bar{j}})$.
 - Se almeno una tra $\bar{\varphi}_1$ e $\bar{\varphi}_2$ usa $<$, allora $\psi_{\bar{i},\bar{j}}$ usa $<$, per densità dell'ordinamento scelgo x_k compreso strettamente tra $v(v_{\bar{i}})$ e $v(w_{\bar{j}})$.

□

Esercizio 3.9. La struttura $(D; <)$ è un ordine totale discreto tale che⁹ $<$ è una relazione d'ordine totale su D tale che ogni elemento di D ha un predecessore ed un successore. $\text{Th}(D; <)$ non elimina i quantificatori, mentre $\text{Th}(D; <, s)$ sì.

Soluzione. Il caso di $\text{Th}(D; <, s)$ è identico a quello della dimostrazione precedente, con l'unica osservazione da fare, che nel caso in cui una tra $\overline{\varphi}_1$ o $\overline{\varphi}_2$ usi $<$, allora posso scegliere x_k come successore di $v(v_i)$ o come predecessore di $v(w_j)$.

Per quanto riguarda $\text{Th}(D; <)$, consideriamo la formula $\varphi := x < y \wedge \neg(\exists z x < z \wedge z < y)$ che esprime che y è il successore immediato di x . Supponiamo per assurdo che esista una formula ψ senza quantificatori, $\text{Th}(D; <)$ -equivalente a φ ; tale formula è necessariamente una combinazione booleana di formule atomiche del tipo $x < y$, $x = y$, \top e \perp . Consideriamo $N := M \times M$ con l'ordinamento lessicografico, allora $N \models \text{Th}(D; <)$. A questo punto consideriamo la coppia $(a, s(a)) \in M^2$, per cui si ha che tale formula è soddisfatta in M e la coppia $((m, a), (m, s(a))) \in N^2$, per un qualsiasi $m \in M$, che rende falsa φ in N . Osserviamo che queste due coppie sono indistinguibili dalle formule atomiche enunciate sopra, in quanto $t_1 < t_2$ in M se e solo se $(m, t_1) < (m, t_2)$ in N , e $t_1 = t_2$ in M se e solo se $(m, t_1) = (m, t_2)$ in N . Segue quindi che ψ ha necessariamente sempre lo stesso valore di verità in M e in N . \square

§3.3 Ultraprodotti

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

Teorema 3.10 (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 3.11 (Ax-Grothendieck)

Sia $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ una funzione polinomiale iniettiva, allora f è surgettiva.

Bando alle ciance.

Definizione 3.12 (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) $A \in F \wedge A \subseteq B \implies B \in F$;

⁹Typo di Mamino.

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

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

$$\forall A, B \in \mathcal{P}(I) \quad A \cup B = I \rightarrow A \in U \vee B \in U$$

Esempio 3.13 (Ultrafiltro principale)

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 3.14 (Filtro dei cofiniti)

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 3.15 (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. Si verifica usando Zorn, e passando per il fatto che l'unione di filtri in catena è un filtro, che ogni filtro F è contenuto in un filtro massimale U . Se questo filtro U non fosse un ultrafiltro, esisterebbero $A, B \subseteq I$ tali che $A \cup B = I$ ma $A, B \notin U$. Consideriamo:

$$G := \{X \subseteq I \mid X \cup B \in U\}$$

è facile verificare che G è un filtro. Inoltre $F \subsetneq G$ perché $A \in F$, ma $A \notin U$. Questo contraddice la massimalità di U , dato che $F \subseteq U \subsetneq G$. \square

Esempio 3.16 (Esistenza di un ultrafiltro non principale)

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 generato da $x \in I$, allora avremmo per definizione di filtro principale che $\{x\} \in U$, allo stesso tempo $I \setminus \{x\}$ appartiene al filtro dei cofiniti (avendo complemento finito), e quindi anche a U . Segue quindi che $\emptyset = \{x\} \cap (I \setminus \{x\}) \in U \not\subseteq U$.

Definizione 3.17 (Ultraprodotto). Fissiamo un linguaggio $L = (R, F)$ e una famiglia di L -strutture $M_i = (D_i; \dots)$ indicizzata da $i \in I$. Sia U un ultrafiltro su $\mathcal{P}(I)$, definiamo l'**ultraprodotto** $\prod_{i \in I} M_i / U$ come la L -struttura avente per dominio $\prod_{i \in I} D_i$ modulo la relazione di equivalenza \sim_U definita come:

$$a, b \in \prod_{i \in I} D_i \quad a \sim_U b \stackrel{\text{def}}{\iff} \{i \in I : a_i = b_i\} \in U$$

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

$$([a_1], \dots, [a_k]) \in r/U \stackrel{\text{def}}{\iff} \{i \in I | (a_{1i}, \dots, a_{ki}) \in r_{M_i}\} \in U^{10}$$

L'interpretazione di $f \in F$ in questo dominio è:

$$f/U([a_1], \dots, [a_k]) \stackrel{\text{def}}{=} [(f_{M_i}(a_{1i}, \dots, a_{ki}))_{i \in I}]^{11}$$

Osservazione 3.18 (Buona definizione dell'ultraprodotto) — Occorre verificare che la definizione è ben posta^a, ossia che se $\underline{a}_1 \sim_U \underline{a}'_1, \dots, \underline{a}_k \sim_U \underline{a}'_k$ allora:

$$([a_1], \dots, [a_k]) \in r/U \iff ([a'_1], \dots, [a'_k]) \in r/U$$

infatti: per ipotesi $([a_1], \dots, [a_k]) \in r/U$ significa che $A = \{i \in I | (a_{1i}, \dots, a_{ki}) \in r_{M_i}\} \in U$. Ora, sempre per ipotesi, sia $B_\iota := \{i \in I | a_{\iota i} = a'_{\iota i}\} \in U, \forall \iota = 1, \dots, k$. Posso quindi considerare:

$$C := A \cap B_1 \cap \dots \cap B_k \in U$$

e, se, $i \in C$, vale che $(a_{1i}, \dots, a_{ki}) = (a'_{1i}, \dots, a'_{ki}) \in r_{M_i}$, di conseguenza, detto $D := \{i \in I | (a'_{1i}, \dots, a'_{ki}) \in r_{M_i}\}$, si ha che $C \subseteq D$, e quindi $D \in U$. Ovviamente la direzione opposta è simmetrica.

Si procede similmente per la buona definizione dell'interpretazione delle funzioni: similmente a prima, sia $B_\iota := \{i \in I | a_{\iota i} = a'_{\iota i}\} \in U, \forall \iota = 1, \dots, k$, e sia $B := B_1 \cap \dots \cap B_k \in U$, allora, se $i \in B$, vale che $(a_{1i}, \dots, a_{ki}) = (a'_{1i}, \dots, a'_{ki})$, e quindi $f_{M_i}(a_{1i}, \dots, a_{ki}) = f_{M_i}(a'_{1i}, \dots, a'_{ki})$. Di conseguenza, $\{i \in I | f_{M_i}(a_{1i}, \dots, a_{ki}) = f_{M_i}(a'_{1i}, \dots, a'_{ki})\} \supseteq B$ per cui appartiene all'ultrafiltro.

^aPer rendere più leggera la trattazione ho deciso di usare la notazione vettoriale per le sequenze di elementi.

Teorema 3.19 (Teorema di Łoś)

$$\prod_{i \in I} M_i/U \models \varphi([a_1], \dots, [a_k]) \iff \{i \in I | M_i \models \varphi(a_{1i}, \dots, a_{ki})\} \in U$$

Notazione 3.20 (Notazione abbreviata per valutazione di formule) — Se $\text{vl}(\varphi) \subseteq \{v_1, \dots, v_k\}$ allora sappiamo che la validità di φ in M dipende solo dai valori assegnati dalla valutazione delle variabili v_1, \dots, v_k . Possiamo indicare questa situazione denotando φ con $\varphi(v_1, \dots, v_k)$ e scrivendo $M \models \varphi(a_1, \dots, a_k)$, con a_1, \dots, a_k nel dominio, per dire che vale $M \models \{v\}\varphi$ a patto che $v(v_1) = a_1, \dots, v(v_k) = a_k$.

¹⁰Cioè la maggioranza degli indici i (indici corrispondenti alle componenti delle varie classi di I -uple, $[a_j]$), secondo l'ultrafiltro, soddisfa la relazione $r \in R$.

¹¹Cioè la funzione f applicata alle componenti delle varie classi di I -uple, $[a_j]$, calcolata indice per indice.

Corollario 3.21 (Un ultraprodotto di modelli è un modello della stessa teoria)

Se per ogni $i \in I$, vale $M_i \models \varphi$, allora $\prod_{i \in I} M_i / U \models \varphi$. Quindi un ultraprodotto di modelli di una teoria T è anch'esso un modello di T . Inoltre, per **U non principale**, se per ogni $i \in I$, **eccetto al più un numero finito di indici**, vale $M_i \models \varphi$, allora $\prod_{i \in I} M_i / U \models \varphi$.

Prima di dare la dimostrazione (un po' noiosa) del teorema di Łoś, vediamo qualche applicazione.

Esempio 3.22 (Modelli non standard di \mathbb{N})

Esistono modelli non isomorfi a \mathbb{N} della teoria $\text{Th}(\mathbb{N}; 0, +, \cdot, s)$, quindi, a fortiori, anche di PA .

Dimostrazione. Sia U un ultrafiltro non principale su $\mathcal{P}(\mathbb{N})$. Consideriamo ${}^*\mathbb{N} = \prod_{i \in \mathbb{N}} \mathbb{N} / U$ ossia l'ultraprodotto (**ultrapotenza**) di una quantità numerabile di copie della medesima struttura $(\mathbb{N}; 0, +, \cdot, s)$. Gli elementi di ${}^*\mathbb{N}$ sono classi di equivalenza di elementi di $\prod_{i \in \mathbb{N}} \mathbb{N}$, cioè di sequenze di numeri naturali. Per il corollario precedente ${}^*\mathbb{N} \models \text{Th}(\mathbb{N}; 0, +, \cdot, s)$. Supponiamo, per assurdo, che $f : \mathbb{N} \rightarrow {}^*\mathbb{N}$ si un isomorfismo di strutture. Allora $f(0_{\mathbb{N}}) = 0_{{}^*\mathbb{N}} = [c_0]$ (dove con c_x indichiamo la successione costante associata ad $x \in \mathbb{N}$, ed è proprio lei $f(0_{\mathbb{N}})$ in quanto per definizione di morfismo di strutture l'immagine è la successione delle immagini di $0_{\mathbb{N}}$ interpretato in ogni modello), perché un morfismo di strutture deve mandare un simbolo di costante interpretato nella prima struttura nello stesso simbolo, interpretato nella seconda. Applicando la funzione successore accade che:

$$\begin{aligned} f(1) &= f(s_{\mathbb{N}}(0_{\mathbb{N}})) && (\text{definizione}) \\ &= s_{{}^*\mathbb{N}}(f(0_{\mathbb{N}})) && (f \text{ morfismo}) \\ &= s_{{}^*\mathbb{N}}(0_{{}^*\mathbb{N}}) \\ &= s_{{}^*\mathbb{N}}([c_0]) && (\text{osservazione sopra}) \\ &= [(s_{\mathbb{N}}(c_{0i}))_{i \in \mathbb{N}}] && (\text{interpretazione di } s \text{ nell'ultraprodotto}) \\ &= [c_{s(0)}] = [c_1] \end{aligned}$$

A questo punto, per induzione, si verifica che $f(n) = [c_n]$ per ogni $n \in \mathbb{N}$, e per violare l'isomorfismo ci basta trovare la classe di una sequenza in ${}^*\mathbb{N}$ che non sia di questo tipo. Considero $[(\sigma(i))_{i \in \mathbb{N}}]$, con $\sigma(i) = i$, in tal caso si ha che:

$$\{i \in \mathbb{N} : \sigma(i) = c_n(i)\} = \{n\}$$

che non può stare nell'ultrafiltro, in quanto lo abbiamo scelto non principale. \square

Esercizio 3.23 (Campi reali chiusi non archimedei). Esistono altri modelli, oltre ad \mathbb{R} , di $\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$ che non soddisfano l'assioma di Archimede:¹²

$$\forall a, b \in M \ (0_M <_M a \wedge 0_M <_M b) \implies \exists n \in \mathbb{N} \ a < \underbrace{b + \dots + b}_{n \text{ volte}}$$

¹²**Hint:** Considera un ultrafiltro non principale su $\mathcal{P}(\mathbb{N})$ e l'ultraprodotto di una quantità numerabile di copie di \mathbb{R} , ${}^*\mathbb{R} = \prod_{i \in \mathbb{N}} \mathbb{R} / U$.

Soluzione. Considero U un ultrafiltro non principale su $\mathcal{P}(\mathbb{N})$ e l'ultrapotenza ${}^*\mathbb{R} = \prod_{i \in \mathbb{N}} \mathbb{R} / U$. Chiaramente, per il corollario di Łoś, ${}^*\mathbb{R} \models \text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$. Considero ora gli elementi $[\underline{c}_1]$ e $[(i)_{i \in \mathbb{N}}]$ in ${}^*\mathbb{R}$. In primis si ha che $0_{{}^*\mathbb{R}} <_{{}^*\mathbb{R}} [\underline{c}_1]$ perché per tutti gli indici l'entrata è positiva; stesso discorso per $0_{{}^*\mathbb{R}} <_{{}^*\mathbb{R}} [(i)_{i \in \mathbb{N}}]$ (tranne per il primo indice, ma questo non conta perché l'ultrafiltro è non principale). A questo punto osservo ancora che $[\underline{c}_1] <_{{}^*\mathbb{R}} [(i)_{i \in \mathbb{N}}]$, perché questa cosa non vale solo per i primi due indici, e per concludere sul fatto che l'assioma di Archimede non vale in ${}^*\mathbb{R}$, suppongo per assurdo che esista un $n \in \mathbb{N}$ tale che:

$$[(i)_{i \in \mathbb{N}}] <_{{}^*\mathbb{R}} \underbrace{[\underline{c}_1] + {}^*\mathbb{R} \dots + {}^*\mathbb{R} [\underline{c}_1]}_{n \text{ volte}} = [c_n]^{13}$$

Ma questo non è possibile perché l'insieme degli indici per cui vale questa diseguaglianza è $\{0, 1, \dots, n\}$ che è finito e quindi non appartiene all'ultrafiltro. \square

§3.4 Teorema di Ax-Grothendieck

Per dimostrare il teorema di Ax-Grothendieck, ci serviranno i seguenti fatti.

Fatto 3.24 (Classificazione dei campi algebricamente chiusi). Due campi algebricamente chiusi della medesima caratteristica sono isomorfi se e solo se hanno basi di trascendenza equipotenti. Di conseguenza ogni campo algebricamente chiuso di caratteristica 0, avente cardinalità 2^{\aleph_0} è isomorfo a \mathbb{C} .

Fatto 3.25 (Chiusura algebrica di un campo finito). La chiusura algebrica $\overline{\mathbb{F}_p}$ di \mathbb{F}_p è $\bigcup_{k \in \mathbb{N}} \mathbb{F}_{p^{k!}}$.

Lemma 3.26 (\mathbb{C} come ultraprodotto di chiusure algebriche di campi finiti)

Sia U un ultrafiltro non principale su $\mathcal{P}(\mathbb{N})$, e sia p_i il i -esimo numero primo. Allora $(\mathbb{C}; 0, 1, +, \cdot)$ è isomorfo a $F = \prod_{i \in \mathbb{N}} \overline{\mathbb{F}_{p_i}} / U$.

Dimostrazione. Il fatto che F sia un campo algebricamente chiuso è esprimibile con una formula nel linguaggio dei campi. Infatti gli assiomi dei campi sono ben noti, inoltre la caratteristica di essere algebricamente chiuso si esprime mediante lo schema di assiomi:

$$\varphi_n = \forall y_0, \dots, y_{n-1} \exists x \ x^n + y_{n-1}x^{n-1} + \dots + y_1x + y_0 = 0 \quad n \in \mathbb{N}$$

Dove x^n è un'abbreviazione per $x \cdot x \dots x$ n volte. Quindi, siccome i fattori dell'ultraprodotto sono campi algebricamente chiusi, anche F lo è. Inoltre F ha caratteristica 0, infatti, fissato un primo p_k :

$$\{i \in \mathbb{N} : \overline{\mathbb{F}_{p_i}} \models \underbrace{1 + \dots + 1 = 0}_{p_k \text{ volte}}\} = \{i \in \mathbb{N} : i = k\} = \{k\} \notin U^{14}$$

per cui F è un campo tale che $F \not\models \underbrace{1 + \dots + 1 = 0}_{p_k \text{ volte}}$, per ogni $k \in \mathbb{N}$; quindi F ha necessariamente caratteristica 0.

¹³La somma nell'ultraprodotto si calcola componente per componente, per definizione di interpretazione dei simboli di funzione.

¹⁴Osservare che se un singoletto appartenesse a un ultrafiltro, allora l'ultrafiltro sarebbe principale.

Resta da dimostrare che $|F| = 2^{\aleph_0}$. Intanto è banale notare che $|F| \leq |\prod_{i \in \mathbb{N}} \overline{\mathbb{F}_{p_i}}| = 2^{\aleph_0}$; per la disegualanza opposta immersiamo i numeri $0, 1, 2, \dots, i$ in $\overline{\mathbb{F}_{p_i}}$ nel modo naturale $i \mapsto 1 + \dots + 1$ i volte, questa mappa è iniettiva perché $i < p_i$. Definiamo ora la funzione:

$$f : [0, 1] \rightarrow F : x \mapsto [(\lfloor x \cdot i \rfloor)_{i \in \mathbb{N}}]$$

e verifico che è iniettiva per concludere. Osserviamo infatti che, per $x < y$, allora esiste $i \in \mathbb{N}$ tale che $\frac{1}{y-x} \leq i$, da cui si ricava $x \cdot i + 1 \leq y \cdot i$, quindi $\lfloor x \cdot i \rfloor + 1 \leq \lfloor y \cdot i \rfloor$, ossia $\lfloor x \cdot i \rfloor < \lfloor y \cdot i \rfloor$; ne segue che:

$$\{i \in \mathbb{N} : \lfloor x \cdot i \rfloor \neq \lfloor y \cdot i \rfloor\} \supseteq \left\{ i \in \mathbb{N} : \frac{1}{y-x} \leq i \right\} \in U$$

e l'insieme a destra è cofinito in quanto il suo complementare è finito (e contiene tutti gli interi strettamente minori di $\frac{1}{y-x}$). \square

Passiamo ora alla dimostrazione del teorema di Ax-Grothendieck.

Dimostrazione. L'idea è che, se \mathbb{C} fosse un insieme finito, l'enunciato sarebbe semplicemente il principio dei cassetti applicato alla funzione f . Ora, \mathbb{C} non è finito, ma dimostreremo che, poiché l'enunciato vale per tutti i campi finiti, allora vale anche per \mathbb{C} . Sia F il campo costruito dal lemma precedente. Fissiamo un grado d arbitrario. Ci basta dimostrare che se $f : F^n \rightarrow F^n$ è una funzione polinomiale iniettiva, allora è surgettiva. Questo enunciato è esprimibile mediante la formula nel linguaggio dei campi: $\varphi_d = \forall z_1, \dots, z_N \text{ Inj}(z_1, \dots, z_N) \rightarrow \text{Surj}(z_1, \dots, z_N)$, con $N = n \cdot \binom{d+n}{n}$ ¹⁵, dove Inj e Surj sono formule che esprimono il fatto che la funzione polinomiale avente coefficienti z_1, \dots, z_N sia rispettivamente iniettiva e surgettiva. In particolare $\text{Surj}(z_1, \dots, z_N)$ avrà la forma:

$$\begin{aligned} \forall y_1, \dots, y_n \exists x_1, \dots, x_n \text{ tali che} \\ y_1 = p_1(z_1, \dots, z_N, x_1, \dots, x_n) \wedge \dots \wedge y_n = p_n(z_1, \dots, z_N, x_1, \dots, x_n) \end{aligned}$$

o in notazione vettoriale, con $\underline{z} = z_1, \dots, z_N$ (coefficienti), $\underline{y} = y_1, \dots, y_n$ (valori che voglio ottenere) e $\underline{x} = x_1, \dots, x_n$ (valori che sto sostituendo alle indeterminate), come $\text{Surj}(\underline{z}) = \forall \underline{y} \exists \underline{x} \underline{y} = \underline{p}(\underline{z}, \underline{x})$.

Per la costruzione di F , per il teorema di Łoś, è sufficiente dimostrare che $\overline{\mathbb{F}_{p_i}} \models \varphi_d$. Fissiamo un p_i primo, $\underline{c} = c_1, \dots, c_N \in \overline{\mathbb{F}_{p_i}}$ e supponiamo che $\overline{\mathbb{F}_{p_i}} \models \text{Inj}(c_1, \dots, c_N)$. Voglio dimostrare che $\overline{\mathbb{F}_{p_i}} \models \forall \underline{y} \exists \underline{x} \underline{y} = \underline{p}(\underline{c}, \underline{x})$.

Fissiamo $\underline{y} = y_1, \dots, y_n \in \overline{\mathbb{F}_{p_i}}$ e cerchiamo $\underline{x} = x_1, \dots, x_n \in \overline{\mathbb{F}_{p_i}}$ tali che risolvano l'equazione $\underline{y} = \underline{p}(\underline{c}, \underline{x})$. Ora sfruttiamo il fatto che $\overline{\mathbb{F}_i} = \bigcup_{j \in \mathbb{N}} \mathbb{F}_{p_i^{j!}}$ per dire che esiste $j_0 \in \mathbb{N}$ tale che $c_1, \dots, c_N, y_1, \dots, y_n \in \mathbb{F}_{p_i^{j_0!}}$. La funzione $\underline{x} \mapsto \underline{p}(\underline{c}, \underline{x})$, definita da $\overline{\mathbb{F}_{p_i}}^n \rightarrow \overline{\mathbb{F}_{p_i}}^n$ si restringe a una funzione iniettiva $g : \mathbb{F}_{p_i^{j_0!}}^n \rightarrow \mathbb{F}_{p_i^{j_0!}}^n$ (l'insieme di arrivo è sempre lo stesso perché sto facendo operazioni nel campo). Sostengo che l' \underline{x} cercato è in $\mathbb{F}_{p_i^{j_0!}}^n$, infatti, siccome $\mathbb{F}_{p_i^{j_0!}}^n$ è finito, allora g è anche surgettiva per il principio dei cassetti, per cui esiste $\underline{x} = x_1, \dots, x_n \in \mathbb{F}_{p_i^{j_0!}}^n$ tale che $\underline{y} = g(\underline{x}) = \underline{p}(\underline{c}, \underline{x})$. \square

Esercizio 3.27 (Controesempio alla freccia inversa di Ax-Grothendieck). L'implicazione contraria (surgettiva \rightarrow iniettiva) è ovviamente falsa. Perché la dimostrazione, in questo verso, non funziona?

¹⁵ $\binom{d+n}{n}$ è il numero di monomi in n variabili di grado totale al più d , in una singola componente.

Soluzione. Chiaramente c'è un controesempio banale alla freccia opposta che è $p(x) = x^2$, come funzione polinomiale da \mathbb{C} in \mathbb{C} .

La dimostrazione sopra non funziona perché non è detto che la restrizione di $f : \overline{\mathbb{F}_{p_i}} \rightarrow \overline{\mathbb{F}_{p_i}}$ a una funzione da $\mathbb{F}_{p_i^{j_0}}$ in se stesso sia surgettiva. \square

§3.5 Compattezza semantica e teorema di Łoś

Procediamo ora alla dimostrazione di compattezza semantica e del teorema di Łoś. Ricordiamo gli enunciati in ambo i casi.

Teorema 3.28 (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 $T' \models \varphi$.

Vediamo prima il caso particolare in cui $|T| = \aleph_0$. Questo caso particolare NON necessita di una dimostrazione separata. Tuttavia questa dimostrazione è più semplice di quella generale, servirà per illustrare meglio l'idea.

Dimostrazione. Sia $T = \{\psi_1, \psi_2, \dots\}$, definiamo $T_i := \{\psi_1, \dots, \psi_i\}$. Ci basta verificare che $\exists i \in \mathbb{N}$ tale che $T_i \models \varphi$, procediamo per assurdo supponendo che $T_i \not\models \varphi$ per ogni $i \in \mathbb{N}$. Abbiamo allora che $\forall i \in \mathbb{N}$ esiste un modello M_i tale che $M_i \models T_i$ ma $M_i \not\models \varphi$, ovvero $M_i \models \neg\varphi$ (questo segue dal fatto che M_i è un controesempio per $T_i \models \varphi$, cioè $M_i \models \neg\varphi$). Sia U un ultrafiltro non principale su $\mathcal{P}(\mathbb{N})$, consideriamo l'ultraprodotto $M = \prod_{i \in \mathbb{N}} M_i / U$, dal corollario al teorema di Łoś, si deduce immediatamente che $M \models \neg\varphi$; se riuscissimo a dimostrare che $M \models T$, avremmo un assurdo, perché per ipotesi $T \models \varphi$.

Per dimostrare che $M \models T$ basta dimostrare che $\forall j \in \mathbb{N}$ vale $M \models \psi_j$, e per il teorema di Łoś questo equivale a dimostrare che $\{i \in \mathbb{N} : M_i \models \psi_j\} \in U$. Segue dall'ipotesi che $M_k \models \psi_j$ per ogni $k \geq j$, quindi:

$$\{i \in \mathbb{N} | M_i \models \psi_j\} \supseteq \{i \in \mathbb{N} | i \geq j\} \in U$$

dove l'ultima appartenenza segue dal fatto che l'insieme è cofinito e U è un ultrafiltro non principale, quindi non può appartenerci un insieme finito, altrimenti ci apparerebbe un singoletto e l'ultrafiltro sarebbe principale. \square

Vediamo ora il **caso generale** del teorema di compattezza semantica.

Dimostrazione. Sia F il sottoinsieme di $\mathcal{P}(\mathcal{P}^{\text{fin.}}(T))$ definito da:

$$X \in F \iff \exists A \in \mathcal{P}^{\text{fin.}}(T) \{B \in \mathcal{P}^{\text{fin.}}(T) | A \subseteq B\} \subseteq X$$

verifichiamo che F è un filtro su $\mathcal{P}(\mathcal{P}^{\text{fin.}}(T))$. Infatti:

- $\emptyset \notin F$ perché il vuoto ha come sottoinsieme solo se stesso, mentre c'è almeno un sottoinsieme non banale finito di T che contiene il vuoto, per cui si ha un assurdo; invece $\mathcal{P}^{\text{fin.}}(T) \in F$ per un qualsiasi $A \in \mathcal{P}^{\text{fin.}}(T)$;
- se $X, Y \in F$ allora esistono $A, B \in \mathcal{P}^{\text{fin.}}(T)$ tali che $\{C \in \mathcal{P}^{\text{fin.}}(T) | A \subseteq C\} \subseteq X$ e $\{D \in \mathcal{P}^{\text{fin.}}(T) | B \subseteq D\} \subseteq Y$; quindi, se $E = A \cup B$, si ha che $\{G \in \mathcal{P}^{\text{fin.}}(T) | E \subseteq G\} \subseteq X \cap Y$, per cui $X \cap Y \in F$;

- se $X \in F$ e $X \subseteq Y \subseteq \mathcal{P}^{\text{fin.}}(T)$, allora esiste $A \in \mathcal{P}^{\text{fin.}}(T)$ tale che $\{B \in \mathcal{P}^{\text{fin.}}(T) | A \subseteq B\} \subseteq X$, per cui $\{B \in \mathcal{P}^{\text{fin.}}(T) | A \subseteq B\} \subseteq Y$, e quindi $Y \in F$.

Sia ora U un ultrafiltro che estende F , supponiamo per assurdo che $\forall T' \in \mathcal{P}^{\text{fin.}}(T)$ si abbia $T' \not\models \varphi$, per cui $\forall T' \in \mathcal{P}^{\text{fin.}}(T)$ esiste $M_{T'}$ tale che $M_{T'} \models T'$ e $M_{T'} \not\models \varphi$, e consideriamo $M := \prod_{T' \in \mathcal{P}^{\text{fin.}}(T)} M_{T'}/_U$. Per il corollario al teorema di Łoś, si ha che $M \models \neg\varphi$. Se verifichiamo che $M \models T$ otteniamo un assurdo; ora $M \models T \iff \forall \psi \in T M \models \psi$, inoltre per il teorema di Łoś ciò equivale a $\{T' \in \mathcal{P}^{\text{fin.}}(T) | M_{T'} \models \psi\} \in U$. Usando la definizione sopra con $A = \{\psi\}$ si ottiene che $\{B \in \mathcal{P}^{\text{fin.}}(T) | A = \{\psi\} \subseteq B\} \subseteq \{T' \in \mathcal{P}^{\text{fin.}}(T) | M_{T'} \models \psi\}$ (il contenimento è ovvio perché se $B \in \text{LHS}$, allora $\psi \in B$, per cui $M_B \models B$, ed in particolare $M_B \models \psi$), per cui $\{T' \in \mathcal{P}^{\text{fin.}}(T) | M_{T'} \models \psi\} \in F \subseteq U$ e si conclude. \square

Concludiamo infine questa sezione con la dimostrazione del teorema di Łoś di cui ricordiamo l'enunciato.

Teorema 3.29 (Teorema di Łoś)

$$\prod_{i \in I} M_i /_U \models \varphi([a_1], \dots, [a_k]) \iff \{i \in I | M_i \models \varphi(a_{1i}, \dots, a_{ki})\} \in U^a$$

^aSia per l'enunciato che per tutta la dimostrazione stiamo usando la notazione compatta per la valutazione delle variabili nelle formule.

Dimostrazione. Siccome, a meno di equivalenza logica, ogni formula può essere scritta usando solamente \exists, \neg, \wedge , possiamo supporre che φ non contenga il quantificatore \forall né altri connettivi salvo \neg e \wedge . Procediamo dunque per induzione strutturale.

Se $\varphi(\underline{x}_1, \dots, \underline{x}_k) = \neg\psi(\underline{x}_1, \dots, \underline{x}_k)$, allora:

$$\begin{aligned} \prod_{i \in I} M_i /_U \models \neg\psi([a_1], \dots, [a_k]) &\iff \neg\left(\prod_{i \in I} M_i /_U \models \psi([a_1], \dots, [a_k])\right) && (\text{Tarski}) \\ &\iff \{i \in I | M_i \models \psi(a_{1i}, \dots, a_{ki})\} \notin U && (\text{hp. ind.}) \\ &\iff \{i \in I | \neg(M_i \models \psi(a_{1i}, \dots, a_{ki}))\} \in U && (\text{def. ultraf.}) \\ &\iff \{i \in I | M_i \models \neg\psi(a_{1i}, \dots, a_{ki})\} \in U && (\text{Tarski}) \end{aligned}$$

Se $\varphi(\underline{x}_1, \dots, \underline{x}_k) = \psi_1(\underline{x}_1, \dots, \underline{x}_k) \wedge \psi_2(\underline{x}_1, \dots, \underline{x}_k)$, allora similmente:

$$\begin{aligned} \prod_{i \in I} M_i /_U \models (\psi_1([a_1], \dots, [a_k]) \wedge \psi_2([a_1], \dots, [a_k])) \\ \iff \prod_{i \in I} M_i /_U \models \psi_1([a_1], \dots, [a_k]) \wedge \prod_{i \in I} M_i /_U \models \psi_2([a_1], \dots, [a_k]) && (\text{Tarski}) \\ \iff \{i \in I | M_i \models \psi_1(a_{1i}, \dots, a_{ki})\} \in U \wedge \{i \in I | M_i \models \psi_2(a_{1i}, \dots, a_{ki})\} \in U && (\text{hp. ind.}) \\ \iff \{i \in I | M_i \models \psi_1(a_{1i}, \dots, a_{ki})\} \cap \{i \in I | M_i \models \psi_2(a_{1i}, \dots, a_{ki})\} \in U && (*) \\ \iff \{i \in I | M_i \models \psi_1(a_{1i}, \dots, a_{ki}) \wedge M_i \models \psi_2(a_{1i}, \dots, a_{ki})\} \in U && (\text{insiemi}) \\ \iff \{i \in I | M_i \models (\psi_1(a_{1i}, \dots, a_{ki}) \wedge \psi_2(a_{1i}, \dots, a_{ki}))\} \in U && (\text{Tarski}) \end{aligned}$$

dove in $(*)$ l'implicazione dal basso verso l'alto è la proprietà 2. della definizione di filtro (chiusura per sovrainsieme), mentre l'implicazione dall'alto verso il basso è la proprietà 3. della definizione di filtro (chiusura per intersezione finita)¹⁶.

¹⁶Notare come abbiano appena dimostrato che dato F filtro, allora $A \cap B \in F \iff A, B \in F$.

Se $\varphi(\underline{x}_1, \dots, \underline{x}_k) = \exists \underline{y} \psi(\underline{x}_1, \dots, \underline{x}_k, \underline{y})$, allora abbiamo ancora che:

$$\begin{aligned} \prod_{i \in I} M_i / U &\models \exists \underline{y} \psi([\underline{a}_1], \dots, [\underline{a}_k], \underline{y}) \\ \iff \exists \underline{b} \in \prod_{i \in I} M_i / U \quad \prod_{i \in I} M_i / U &\models \psi([\underline{a}_1], \dots, [\underline{a}_k], [\underline{b}]) && \text{(Tarski)} \\ \iff \exists \underline{b} \in \prod_{i \in I} M_i / U \quad (\{i \in I \mid M_i \models \psi(a_{1i}, \dots, a_{ki}, b_i)\} \in U) && \text{(hp. ind.)} \\ \iff \{i \in I \mid \exists b_i \in M_i \text{ t.c. } M_i \models \psi(a_{1i}, \dots, a_{ki}, b_i)\} \in U && (\star\star) \\ \iff \{i \in I \mid M_i \models \exists \underline{y} : \psi(a_{1i}, \dots, a_{ki}, \underline{y})\} \in U && \text{(Tarski)} \end{aligned}$$

dove in $(\star\star)$ l'implicazione dall'alto verso il basso è sempre la proprietà 2. della definizione di filtro (chiusura per sovrainsieme), mentre l'implicazione dal basso verso l'alto è AC in quanto posso costruire \underline{b} scegliendo per ogni $i \in I$ un b_i tale che $M_i \models \psi(a_{1i}, \dots, a_{ki}, b_i)$ (non servono tutti ma solo quelli che funzionano), che esiste per ipotesi.

Per completezza vediamo anche i casi in cui φ è una L -formula atomica.

Se $\varphi(\underline{x}_1, \dots, \underline{x}_k) = \top, \perp$, allora è banale, infatti $\{i \in I \mid M_i \models \top\} = I \in U$, e per definizione di ultrafiltro $\{i \in I \mid M_i \models \perp\} = \emptyset \notin U$.

Se $\varphi(\underline{x}_1, \dots, \underline{x}_k) = r(t_1, \dots, t_m)(x_1, \dots, x_k)$, con r simbolo di relazione di arietà m , allora:

$$\begin{aligned} \prod_{i \in I} M_i / U &\models r([\underline{a}_1], \dots, [\underline{a}_k]) \\ \iff ([\underline{a}_1], \dots, [\underline{a}_k]) &\in r / U && \text{(Tarski)} \\ \iff \{i \in I \mid (a_{1i}, \dots, a_{ki}) \in r_{M_i}\} &\in U && \text{(interpretazione di } r \text{ nell'ultraprodotto)} \\ \iff \{i \in I \mid M_i \models r(a_{1i}, \dots, a_{ki})\} &\in U && \text{(Tarski)} \end{aligned}$$

Se $\varphi(\underline{x}_1, \dots, \underline{x}_k) = (t_1(\underline{x}_1, \dots, \underline{x}_k) = t_2(\underline{x}_1, \dots, \underline{x}_k))$, con t_1, t_2 L -termini, allora:

$$\begin{aligned} \prod_{i \in I} M_i / U &\models t_1([\underline{a}_1], \dots, [\underline{a}_k]) = t_2([\underline{a}_1], \dots, [\underline{a}_k]) \\ \iff [t_1([\underline{a}_1], \dots, [\underline{a}_k])] &= [t_2([\underline{a}_1], \dots, [\underline{a}_k])] && (\star) \\ \iff \{i \in I \mid t_1(a_{1i}, \dots, a_{ki}) = t_2(a_{1i}, \dots, a_{ki})\} &\in U && \text{(def. ultraprodotto)} \\ \iff \{i \in I \mid M_i \models (t_1(a_{1i}, \dots, a_{ki}) = t_2(a_{1i}, \dots, a_{ki}))\} &\in U && \text{(Tarski)} \end{aligned}$$

dove in (\star) stiamo usando la semantica di Tarski ed otteniamo un'uguaglianza di L -termini nel modello (l'ultraprodotto), quindi un'uguaglianza di classi di equivalenza di successioni. \square

§3.6 Applicazioni del teorema di compattezza

In questa sezione mostriamo alcuni esempi ed applicazioni del teorema di compattezza. Le tecniche che useremo costituiscono gli elementi della **teoria dei modelli**, che studia le proprietà e della relazione di conseguenza logica. I risultati principali che dimostreremo sono: i teoremi di Löwenheim-Skolem - ce ne sono due, uno per salire e uno per scendere in cardinalità, però, rozzamente, possiamo scrivere come segue.

Teorema 3.30 (Löwenheim-Skolem - alla buona)

O i modelli di una L -teoria T hanno tutti cardinalità $\leq n$ per qualche $n \in \mathbb{N}$. Oppure, per ogni cardinalità $\kappa \geq |L| + \aleph_0$, la teoria T ha un modello di cardinalità κ .

Da questo risultato segue un criterio che, per esempio, ci permetterà di dare una dimostrazione rapida del fatto che T_{oldse} è completa, oppure di dimostrare che la teoria dei campi algebricamente chiusi di caratteristica 0 è completa. Cominciamo innanzitutto con qualche esempio.

Proposizione 3.31 (Finitamente coerente \implies coerente)

Se ogni sottoteoria finita di una teoria T è coerente - ossia se T è **finitamente coerente** - allora T è coerente.

Dimostrazione. Per la caratterizzazione vista T è coerente se e solo se $T \not\models \perp$, pertanto, se per assurdo $T \models \perp$, per la compattezza semantica esisterebbe un sottoinsieme finito $T' \subseteq T$ tale che $T' \models \perp$, ma per ipotesi T' è coerente, per cui si ha un assurdo. Segue quindi che $T \not\models \perp$, che equivale a dire che T è coerente. \square

Esempio 3.32 (Modelli non standard di $\text{Th}(\mathbb{N}; 0, 1, +, \cdot, s)$)

Abbiamo già visto che ci sono modelli non standard di $\text{Th}(\mathbb{N}; 0, 1, +, \cdot, s)$. Dimostriamo questo fatto per compattezza semantica.

Dimostrazione. Sia $L_c = \{\underline{0, 1, +, \cdot, s}\} \cup \{c_n | n \in \mathbb{N}\} = L_{\text{ar}} \cup \{\underline{c}\}$, il linguaggio dell'aritmetica **espanso** con un nuovo simbolo di costante \underline{c} . Consideriamo la L_c -teoria:

$$T = \underbrace{\text{Th}(\mathbb{N}; 0, 1, +, \cdot, s)}_{L_{\text{ar}}} \cup \underbrace{\{\exists x c = s(x), \exists x c = s(s(x)), \dots\}}_{\text{nuovi assiomi}}$$

Questa teoria è finitamente coerente perché, data $T' \subseteq T$ finita, T' non può che contenere un numero finito di nuovi assiomi, quindi \mathbb{N} , interpretando \underline{c} come un numero abbastanza grande, è un modello di T' . Per la proposizione precedente, T è coerente, quindi ha un modello M .

Ora M è una L_c -struttura e $M \models \text{Th}(\mathbb{N}; 0, 1, +, \cdot, s)$. Consideriamo il **ridotto** $M|_{L_{\text{ar}}}$ di M a L_{ar} - ossia, se $M = (D; i)$, la L_{ar} -struttura $M|_{L_{\text{ar}}} = (D; i|_{L_{\text{ar}}})$. È chiaro che $M|_{L_{\text{ar}}} \models \text{Th}(\mathbb{N}; L_{\text{ar}})$, inoltre $M|_{L_{\text{ar}}}$ non è isomorfa a \mathbb{N} perché l'elemento c_M , che appartiene al suo dominio, ha una catena infinita di predecessori, per cui un eventuale isomorfismo di L_{ar} -struttura con \mathbb{N} genererebbe un'infinita catena di predecessori in \mathbb{N} . \square

Esercizio 3.33 (Modelli non standard di $\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$). Allo stesso modo si può dimostrare anche che esistono modelli non standard di $\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$.¹⁷

Soluzione. Sia $L_c = \underbrace{\{\underline{0, 1, +, \cdot, <}\}}_{L_{\text{co}}} \cup \{\underline{c}\}$, il linguaggio dei campi ordinati espanso con una nuova costante \underline{c} . Consideriamo la L_c -teoria:

$$T = \underbrace{\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)}_{L_{\text{co}}} \cup \underbrace{\{\underline{1 < c}, \underline{1 + 1 < c}, \underline{1 + 1 + 1 < c}, \dots\}}_{\text{nuovi assiomi}}$$

Questa teoria è finitamente coerente perché, data $T' \subseteq T$ finita, T' non può che contenere un numero finito di nuovi assiomi, quindi \mathbb{R} , interpretando \underline{c} come un numero abbastanza

¹⁷Hint: Basata aggiungere una costante \underline{c} e gli assiomi $\{1 < c, 1 + 1 < c, 1 + 1 + 1 < c, \dots\}$.

grande, è un modello di T' . Per la proposizione precedente, T è coerente, quindi ha un modello M .

Ora M è una L_c -struttura e $M \models \text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$. Consideriamo il ridotto $M|_{L_{co}}$ di M a L_{co} - ossia, se $M = (D; i)$, la L_{co} -struttura $M|_{L_{co}} = (D; i|_{L_{co}})$. È chiaro che $M|_{L_{co}} \models \text{Th}(\mathbb{R}; L_{co})$, inoltre $M|_{L_{co}}$ non è isomorfa a \mathbb{R} perché l'elemento $c_{M|_{L_{co}}}$, che appartiene al suo dominio, è maggiore di ogni numero reale, per cui un eventuale isomorfismo di L_{co} -struttura con \mathbb{R} genererebbe un numero reale più grande di ogni numero reale in \mathbb{R} , il che violerebbe la proprietà Archimedea ζ . Per cui $M|_{L_{co}}$ è un modello di $\text{Th}(\mathbb{R}; 0, 1, +, \cdot, <)$ che non rispetta la proprietà Archimedea. \square

Definizione 3.34 (Assiomatizzabilità e finita assiomatizzabilità). Diciamo che una classe C di L -strutture è **assiomatizzabile** se c'è una L -teoria T tale che una L -struttura appartiene a C se e solo se è un modello di T . Se c'è una T finita siffatta, allora C è **finitamente assiomatizzabile**.

Esempio 3.35 (Assiomatizzabilità della classe dei buoni ordini)

La classe dei buoni ordini, nel linguaggio $L = \{\langle\}$ non è assiomatizzabile.

Dimostrazione. Supponiamo, per assurdo, che $M \models T$ se e solo se M è un buon ordine. Allora:

$$T' = T \cup \{c_2 < c_1, c_3 < c_2, c_4 < c_3, \dots\}$$

è una teoria coerente nel linguaggio:

$$L' = L \cup \{c_1, c_2, c_3, \dots\}$$

infatti, data $T'' \subseteq T'$ finita, avremo:

$$T'' \subseteq T \cup \{c_2 < c_1, \dots, c_n < c_{n-1}\}$$

con $n \in \mathbb{N}$, per cui ω con $c_i = n - i$ è un modello di T'' , quindi per la proposizione precedente T' è coerente. Tuttavia, detto M un modello di T' , dovremmo avere che $M|_L \models T$, ma questo contraddice il fatto che le interpretazioni delle costanti c_1, c_2, \dots in M formano una catena discendente infinita. \square

Esercizio 3.36 (Classi di strutture non assiomatizzabili). Le seguenti classi di strutture NON sono assiomatizzabili:

1. insiemi finiti nel linguaggio \emptyset ;
2. grafi connessi - con il linguaggio $L = \{e(\cdot, \cdot)\}$ (simbolo di relazione binaria) - per ogni coppia di vertici è collegata esiste un cammino che li connette;
3. campi di caratteristica finita;
4. *gruppi liberi - esistono un insieme di generatori liberi, ossia senza relazioni non banali tra di loro;
5. *gruppi semplici - ossia senza sottogruppi normali non banali.

Soluzione. Vediamo i vari casi separatamente.

1. Supponiamo, per assurdo, che esista una teoria T che assiomatizza la classe degli insiemi finiti nel linguaggio vuoto $L_\emptyset = \emptyset$. Allora:

$$\begin{aligned} T' = T \cup \{ & \exists x_1 \exists x_2 \neg(x_1 = x_2), \\ & \exists x_1 \exists x_2 \exists x_3 \neg(x_1 = x_2) \wedge \neg(x_1 = x_3) \wedge \neg(x_2 = x_3), \dots \} \end{aligned}$$

è una L -teoria nel linguaggio espanso $L = L_\emptyset \cup \{\textcolor{red}{c}_1, \textcolor{red}{c}_2, \dots\}$. Osserviamo che T' è finitamente coerente perché data una qualsiasi sottoteoria finita questa conterrà un numero finito di nuovi assiomi, per cui un sottoinsieme sufficientemente grande di \mathbb{N} è un modello di tale sottoteoria. Per la proposizione precedente, T' è coerente, quindi ha un modello M . A questo punto consideriamo la struttura ridotta $M|_{L_\emptyset}$, che è un modello di T , ma non è finita perché le interpretazioni delle costanti $\textcolor{red}{c}_1, \textcolor{red}{c}_2, \dots$ in M , che sono infinite e distinte, appartengono al dominio di $M|_{L_\emptyset}$, per cui si ha un assurdo.

2. Supponiamo, per assurdo, che esista una teoria T che assiomatizza la classe dei grafi connessi nel linguaggio $L = \{\textcolor{red}{e}(\cdot, \cdot)\}$.

3. Se per assurdo che esista una teoria T che assiomatizza la classe dei campi di caratteristica finita nel linguaggio dei campi $L = \{\textcolor{red}{0}, \textcolor{red}{1}, +, \cdot\}$, per quanto visto:

$$\mathbb{C} = \prod_{p \text{ primo}} \overline{\mathbb{F}_p}/U$$

con U ultrafiltro non principale. Ora se F è un campo di caratteristica finita se e solo se $F \models T$, per il teorema di Łoś si ha che $\mathbb{C} \models T$, ma questo è assurdo perché \mathbb{C} ha caratteristica 0.

4. Supponiamo per assurdo che esista una teoria T che assiomatizza la classe dei gruppi liberi nel linguaggio dei gruppi $L_{\text{gr}} = \{\cdot, ^{-1}, \textcolor{red}{e}\}$. Allora, preso U ultrafiltro non principale su $\mathcal{P}(\mathbb{N})$, possiamo considerare l'ultraprodotto:

$$G = \prod_{n \in \mathbb{N}} \mathbb{Z}/U$$

che è un gruppo abeliano e libero in quanto $\mathbb{Z} \models T$ per ipotesi (dunque $G \models T$ per il teorema di Łoś). Pertanto l'unica possibilità è che $G \cong \mathbb{Z}$, ma questo è assurdo perché G ha cardinalità 2^{\aleph_0} .

Verifichiamo quest'ultima cosa, naturalmente, $|\prod_{n \in \mathbb{N}} \mathbb{Z}/U| \leq |\prod_{n \in \mathbb{N}} \mathbb{Z}| = 2^{\aleph_0}$; per la disegualanza opposta, consideriamo la funzione:

$$f : [-1, 1] \rightarrow \prod_{n \in \mathbb{N}} \mathbb{Z}/U : x \mapsto [(\lfloor x \cdot i \rfloor)_{i \in \mathbb{N}}]$$

Vediamo che f è iniettiva, siano $x < y$ in $[-1, 1]$, allora vale che $\frac{1}{y-x} \leq i$ definitamente, ma questa disegualanza equivale a dire che $\lfloor x \cdot i \rfloor < \lfloor y \cdot i \rfloor$, per cui:

$$\{i \in \mathbb{N} \mid \lfloor x \cdot i \rfloor \neq \lfloor y \cdot i \rfloor\} \supseteq \left\{ i \in \mathbb{N} \mid i \geq \frac{1}{y-x} \right\} \in U$$

dove l'ultima appartenenza vale perché U è un ultrafiltro non principale e quell'insieme è cofinito.

5. Supponiamo per assurdo che esista una teoria T che assiomatizza la classe dei gruppi semplici nel linguaggio dei gruppi $L_{\text{gr}} = \{\cdot, -, e\}$. Per cui un gruppo G è semplice se e solo se $G \models T$. Osserviamo che (per il teorema di Lagrange) \mathbb{F}_p è semplice per ogni p primo. Consideriamo quindi l'ultraprodotto:

$$\mathbb{C} \cong \prod_{p \text{ primo}} \mathbb{F}_p / U$$

con U ultrafiltro non principale; per il teorema di Łoś si avrebbe che $\mathbb{C} \models T$, ma questo è assurdo perché \mathbb{C} non è semplice in quanto ha sottogruppi normali non banali (ad esempio \mathbb{Z}). □

Esempio 3.37 (Assiomatizzabilità della classe degli insiemi infiniti)

La classe degli insiemi infiniti nel linguaggio vuoto L_\emptyset , è assiomatizzabile ma non finitamente assiomatizzabile.

Dimostrazione. Sia:

$$\begin{aligned} \text{AtLeast}_n &= \exists x_1 \exists x_2 \dots \exists x_n \underbrace{\neg(x_1 = x_2) \wedge \neg(x_1 = x_3) \wedge \dots}_{=} \\ &= \bigwedge_{\substack{i < j \leq n \\ i \neq j}} \neg(x_i = x_j) \end{aligned}$$

e sia $T = \{\text{AtLeast}_1, \text{AtLeast}_2, \dots\}$, è chiaro che tale teoria assiomatizza la classe degli insiemi infiniti (ogni insieme infinito rispetta tutte quelle formule, ed al contempo il soddisfare tutte quelle formule garantisce poter scegliere infiniti elementi distinti). Così abbiamo dimostrato che la teoria degli insiemi infiniti nel linguaggio vuoto è assiomatizzabile.

Supponiamo ora che sia finitamente assiomatizzabile, ossia che esista T' una assiomatizzazione finita. Si ha ovviamente che $T \models T'$, e, come conseguenza, del teorema di compattezza semantica, si ha che $\exists n \in \mathbb{N}$ abbastanza grande tale che $T'' := \{\text{AtLeast}_1, \dots, \text{AtLeast}_n\} \models T'$, per cui anche T'' assiomatizza la classe degli insiemi infiniti nel linguaggio vuoto.

Ora osserviamo che anche T'' caratterizza la classe degli insiemi infiniti, infatti: se M è un insieme infinito, allora, banalmente, $M \models T''$; viceversa, se $M \models T''$, allora $M \models T'$, per cui M è infinito. Siamo dunque arrivati ad un assurdo perché $\{1, \dots, n\} \models T''$. □

Esercizio 3.38 (Classi di strutture assiomatizzabili ma non finitamente assiomatizzabili). Le seguenti classi di strutture sono assiomatizzabili ma non finitamente assiomatizzabili:

1. gruppi/anelli/campi infiniti;
2. campi di caratteristica 0;
3. campi algebricamente chiusi;
4. gruppi abeliani divisibili $\forall n \in \mathbb{N} \forall x \in G \exists y \in G x = \underbrace{y + \dots + y}_{n \text{ volte}}$ ¹⁸

¹⁸Mamino lo ha piazzato nella sezione sbagliata, ma il suo posto è qui.

5. *grafi 3-colorabili - esiste una partizione dei vertici in tre sottoinsiemi, nessuno dei quali contiene due vertici adiacenti.

Soluzione. Vediamo i vari casi separatamente.

1. Nel linguaggio $L_{\text{gr}} = \{\cdot, -, e\}$ dei gruppi, un'assiomatizzazione della classe dei gruppi infiniti è data da:

$$T := \{\text{assiomi dei gruppi}\} \cup \{\text{AtLeast}_1, \text{AtLeast}_2, \dots\}$$

dove AtLeast_n è come definita nell'esempio precedente. Chiaramente G è un gruppo infinito se e solo se $G \models T$. Supponiamo ora per assurdo che esista T' finita che assiomatizza la classe dei gruppi infiniti, allora ovviamente $T' \models T$, e per il teorema di compattezza semantica esiste $n \in \mathbb{N}$ abbastanza grande tale che $T'' := \{\text{assiomi dei gruppi}\} \cup \{\text{AtLeast}_1, \dots, \text{AtLeast}_n\} \models T'$. Osserviamo ora che anche T'' assiomatizza la classe dei gruppi infiniti, infatti: se G è un gruppo infinito, allora banalmente $G \models T''$; viceversa, se $G \models T''$, allora $G \models T'$, per cui G è infinito. Siamo dunque arrivati ad un assurdo perché $\mathbb{Z}/(n) \models T''$.

Il ragionamento funziona uguale per gli anelli e per i campi, con l'unica osservazione che nel caso dei campi scelgo come n un primo p sufficiente grande per poi usare \mathbb{F}_p come controesempio.

2. Consideriamo la formula: $\underbrace{1 + 1 + \dots + 1}_{p \text{ volte}} = 0 =: \text{char}_p$, allora un'assiomatizzazione della classe dei campi di caratteristica 0 nel linguaggio dei campi è data da:

$$T := \{\text{assiomi dei campi}\} \cup \{\neg \text{char}_{p_i} | p_i \text{ primo}, i \in \mathbb{N}\}$$

Chiaramente un campo F ha caratteristica 0 se e solo se $F \models T$. Supponiamo ora per assurdo che esista T' finita che assiomatizza la classe dei campi di caratteristica 0, allora chiaramente $T' \models T$, e per il teorema di compattezza semantica esiste un $n \in \mathbb{N}$ sufficientemente grande tale che $T'' := \{\text{assiomi dei campi}\} \cup \{\neg \text{char}_{p_1}, \dots, \neg \text{char}_{p_n}\} \models T'$. Osserviamo ora che anche T'' assiomatizza la classe dei campi di caratteristica 0, infatti: se F è un campo di caratteristica 0, allora banalmente $F \models T''$; viceversa, se $F \models T''$, allora $F \models T'$, per cui F ha caratteristica 0. Siamo dunque arrivati ad un assurdo perché $\mathbb{F}_{p_{n+1}} \models T''$ ed ha caratteristica p_{n+1} .

3. Nel linguaggio dei campi definiamo, $\forall n \geq 1$, la formula: $\forall y_0, \dots, y_{n-1} \exists x x^n + y_{n-1}x^{n-1} + \dots + y_0 = 0 =: \varphi_n$, che esprime il fatto che ogni polinomio di grado n ha una radice¹⁹. Possiamo quindi assiomatizzare la classe dei campi algebricamente chiusi con la teoria:

$$T := \{\text{assiomi dei campi}\} \cup \{\varphi_n | n \in \mathbb{N}, n \geq 1\}$$

È chiaro che un campo F è algebricamente chiuso se e solo se $F \models T$. Supponiamo ora per assurdo che esista T' finita che assiomatizza la classe dei campi algebricamente chiusi, allora come prima esiste $n \in \mathbb{N}$ sufficientemente grande tale che: $T'' := \{\text{assiomi dei campi}\} \cup \{\varphi_1, \dots, \varphi_n\}$ assiomatizza la classe dei campi algebricamente chiusi. [METTERE IL CONTROESEMPIO CON L'UNIONE DI CAMPI]

¹⁹Naturalmente, essendo in un campo, è sufficiente considerare polinomi monici.

4. Nel linguaggio dei gruppi possiamo fornire un'assiomatizzazione esplicita della classe dei gruppi abeliani divisibili come segue:

$$T := \{\text{assiomi dei gruppi} + \text{abelianità}\} \cup \{\psi_n \mid n \in \mathbb{N}, n \geq 1\}$$

dove ψ_n è la L_{gr} -formula: $\forall x \exists y x = ny = x$, che esprime la divisibilità per n .

5. Sia $L = \{\textcolor{red}{e}(\cdot, \cdot)\}$ il linguaggio dei grafi. Per ogni $n \in \mathbb{N}$ definiamo la L -formula:

$$\chi_n = \exists v_1, \dots, v_n \bigwedge_{1 \leq i < j \leq n} \neg e(v_i, v_j)$$

e sia $T = \{\text{assiomi dei grafi}\} \cup \{\chi_n \mid n \in \mathbb{N}\}$. È chiaro che un grafo G è 3-colorabile se e solo se $G \models T$. Inoltre, per un ragionamento simile a quello fatto per gli insiemi infiniti, si può dimostrare che T non è finitamente assiomatizzabile.

□

Vediamo ora due esempi di applicazione al di fuori della logica matematica.

Definizione 3.39 (Partizione litigiosa). Diciamo che una partizione $V = A \sqcup B$ dei vertici di un grafo $G = (V, E)$ è **litigiosa** se, per ogni vertice $v \in V$, il numero di adiacenti a v che appartiene alla medesima parte di v è \leq del numero di adiacenti a v che appartiene all'altra parte.

Proposizione 3.40

Sia G un grafo **localmente finito** - ossia ogni vertice di G ha un numero finito di vertici adiacenti - allora G ammette una partizione litigiosa.

DA INSERIRE.

□

La seconda applicazione ci dà il pretesto per introdurre le seguenti definizioni, che, in realtà, sono nozioni di base della teoria dei modelli.

Definizione 3.41 (Linguaggio espanso). Sia $M = (D; \dots)$ una L -struttura, $L(M)$ è il linguaggio **L -espanso** con l'aggiunta di una costante $\textcolor{red}{c}_i$ per ogni $i \in D$. Possiamo vedere M come una $L(M)$ -struttura, denotata $\textcolor{blue}{M}_M$, estendendo la funzione interpretazione di M , in modo che interpreti ogni $\textcolor{red}{c}_i$ con i .

Definizione 3.42 (Diagramma elementare). Il **diagramma elementare** di una L -struttura M , denotato con $\text{ED}(M)$ è la $L(M)$ -teoria $\text{Th}(M_M; L(M))$ - ossia l'insieme di tutte le $L(M)$ -formule valide in M_M .

Definizione 3.43 (Diagramma atomico). Il **diagramma atomico** di una L -struttura M , denotato con $\text{diag}(M)$, è l'insieme delle $L(M)$ -formule atomiche o negazioni di atomiche valide in M_M - quindi è il sottoinsieme di $\text{Th}(M_M; L(M))$ costituito solo da formule atomiche e loro negazioni.

Osservazione 3.44 — È chiaro che $\text{diag}(M) \subseteq \text{ED}(M)$.

Definizione 3.45 (Sottostruttura). Sia $N = (D; i)$ una L -struttura e $C \subseteq D$. Se per ogni simbolo di funzione f di L abbiamo che $f_N[C^{\text{ar}(f)}] \subseteq C$ - ossia C è chiuso per l'operazione f - allora $M = (C; i|_C)$, dove $i|_C$ è ottenuta restringendo il dominio di relazioni e funzioni, si dice **sottostruttura** di N , e si denota $M \subseteq N$.

Definizione 3.46 (Sottostruttura elementare). Se $M = (C; i|_C)$ è una sottostruttura di $N = (D; i)$ e per ogni L -formula $\varphi(x_1, \dots, x_k)$ vale:

$$\forall a_1, \dots, a_k \in C \quad M \models \varphi(a_1, \dots, a_k) \iff N \models \varphi(a_1, \dots, a_k)$$

- cioè se la verità di φ non dipende dagli elementi esterni alla sottostruttura - allora M si dice **sottostruttura elementare** di N , e si denota $M \preceq N$.

Definizione 3.47 (Estensione/Estensione elementare). Infine, se $M \subseteq N$ diciamo che è N è un'**estensione** di M ; e analogamente se $M \preceq N$ diciamo che N è una **estensione elementare** di M .

Osservazione 3.48 (Caratterizzazione delle sottostrutture tramite i diagrammi) —

Siano $M = (C; \dots)$ e $N = (D; \dots)$ con $C \subseteq D$. La L -struttura N può essere vista come una $L(M)$ -struttura N_M , interpretando ogni simbolo di costante di $L(M)$ come in M_M . Vale allora che:

- $M \subseteq N$ se e solo se $N_M \models \text{diag}(M)$;
- $M \preceq N$ se e solo se $N_M \models \text{ED}(M)$.

METTERE IN UN'APPENDICE. □

Basta burocrazia, questa è l'applicazione promessa.

Teorema 3.49 (Levi)

Ogni gruppo abeliano senza torsione è ordinabile - ossia esiste una relazione di ordine totale $<$ tale che $\forall a, b, c \in G \quad a < b \implies ac < bc$.

Sfrutteremo la seguente osservazione.

Osservazione 3.50 (Sottostrutture di modelli di teorie universali) — Se T è una **teoria universale** - ossia tutte le $\varphi \in T$ sono della forma $\forall x_1 \dots \forall x_k \psi$, con ψ senza quantificatori - e M è una sottostruttura di un modello di T , allora M è un modello di T .

Vediamo la dimostrazione dell'osservazione in primis.

Esercizio 3.51 (Formule assolute ed universali). Dimostra che le formule senza quantificatori sono **assolute** - valgono nella sottostruttura se e solo se valgono nell'estensione. Le formule **universali** si preservano per sottostrutture. Mentre le formule **esistenziali** si preservano per estensioni.

Veniamo ora alla dimostrazione del teorema di Levi.

Dimostrazione. Sia T_{oag} la teoria dei gruppi abeliani ordinati nel linguaggio $L_{\text{oag}} = \{e, ^{-1}, \cdot, <\}$, che è assiomatizzata da:

$$T_{\text{oag}} = T_{\text{gr. ab.}} \cup T_{\text{ord. tot.}} \cup \{\forall x, y, z \quad x < y \implies xz < yz\}$$

si vede che T_{oag} è una teoria universale. Sia ora G un gruppo abeliano senza torsione, è sufficiente trovare un modello di $T_{\text{oag}} \cup \text{diag}(G)$, infatti, in tal caso, avrei che $G' := G_G$, l'insieme delle interpretazioni delle costanti c_g , per $g \in G$, è una sottostruttura del modello in questione che, come gruppo, è isomorfa a G ; per l'osservazione precedente, si avrebbe allora che $G' \models T_{\text{oag}}$, e quindi via isomorfismo $G \models T_{\text{oag}}$.

Resta da mostrare che $T_{\text{oag}} \cup \text{diag}(G)$ è coerente. Se per assurdo non fosse coerente, ci sarebbe un sottoinsieme finito $Y \subseteq T_{\text{oag}} \cup \text{diag}(G)$ tale che Y è incoerente, allora, se prendo $X := Y \cap \text{diag}(G)$ ho che $T_{\text{oag}} \cup X$ è incoerente (mi basta meno di T_{oag} per arrivare all'incoerenza).

Consideriamo ora il sottogruppo H di G generato dai g corrispondenti alle $c_g \in X$ che compaiono in X^{20} . Questo è un gruppo abeliano finitamente generato (X è finito, quindi anche i simboli di costante di G interpretati che vi appaiono) e senza torsione, per cui, per il teorema di struttura dei moduli finitamente generati su PID, vale che $H \cong \mathbb{Z}^n$ per qualche $n \in \mathbb{N}$. A questo punto, ordinando H tramite l'ordinamento lessicografico, ottengo che $H \models T_{\text{oag}} \cup X$ (soddisfa X perché l'ho definito apposta per interpretare le sue costanti), il che è assurdo. \square

§3.7 Teoremi di Löwenheim-Skolem

Teorema 3.52 (Löwenheim-Skolem verso l'alto - forma debole)

Il teorema di Löwenheim-Skolem verso l'alto dice due cose.

1. Sia T una L -teoria. Supponiamo che, per ogni $n \in \mathbb{N}$, ci sia un modello di T di cardinalità $\geq n$. Allora, per ogni cardinalità κ , c'è un modello di T di cardinalità $\geq \kappa$.
2. Sia M una L -struttura infinita, allora, per ogni cardinalità κ , c'è una estensione elementare N di M avente cardinalità $\geq \kappa$.

Dimostrazione. Per il punto 1. è sufficiente espandere L aggiungendo κ costanti c_i , con $i \in \kappa$. Allora la teoria:

$$T' = T \cup \{\neg c_i = c_j \mid i, j \in \kappa, i \neq j\}$$

è finitamente coerente, infatti, data $T'' \subseteq T'$ finita, T'' contiene un numero finito di nuovi assiomi, per cui esiste un modello di T di cardinalità abbastanza grande che interpreta le costanti c_i in modo distinto (semplicemente per ipotesi). Segue quindi che T' è coerente, per cui ha un modello M , che necessariamente ha cardinalità $\geq \kappa$, perché interpreta le costanti c_i in modo distinto.

Per il punto 2. basta applicare il punto 1. alla teoria $\text{ED}(M)$, che ha un modello, M_M , di cardinalità $\geq |M|$ (perché deve interpretare tutte le costanti distinte). \square

Vorremo ora rimpiazzare $\geq \kappa$ con $= \kappa$. Ci servirà il lemma seguente, che è spesso utile per costruire sottostrutture elementari.

²⁰Typo Mamino.

Lemma 3.53 (Criterio di Tarski-Vaught)

Sia $M \subseteq N$ due L -strutture. Allora $M \preceq N$ se e solo se, per ogni L -formula $\varphi(\textcolor{red}{x}, \textcolor{red}{y}_1, \dots, \textcolor{red}{y}_k)$ vale:

$$\forall b_1, \dots, b_k \in M \ N \models \exists \textcolor{red}{x} \ \varphi(\textcolor{red}{x}, b_1, \dots, b_k) \implies \exists \textcolor{red}{a} \in M \ N \models \varphi(\textcolor{red}{a}, b_1, \dots, b_k)$$

ovvero se una formula con parametri in M è soddisfatta in N , allora è soddisfatta in M da un elemento di M .

Nota 3.54 (Utilità del criterio di Tarski-Vaught) — La condizione del criterio di Tarski-Vaught non menziona $M \models$. Ottimo se stiamo cercando di costruire M , per cui non lo abbiamo ancora fissato.

DA INSERIRE.

□

Teorema 3.55 (Löwenheim-Skolem verso il basso)

Sia N una L -struttura infinita, Sia A un sottoinsieme del dominio di N . Sia infine κ un cardinalità infinita con $|L| + |A| \leq \kappa \leq |N|$. Allora esiste $M \preceq N$ con $|M| = \kappa$ il cui dominio contiene A .

DA INSERIRE.

□

Teorema 3.56 (Löwenheim-Skolem - verso l'alto - forma forte)

Come prima, ma con cardinalità esatta.

1. Sia T una L -teoria. Supponiamo che, per ogni $n \in \mathbb{N}$, ci sia un modello di T di cardinalità $\geq n$. Allora, per ogni cardinalità $\kappa \geq |L| + \aleph_0$, c'è un modello di T di cardinalità $= \kappa$.
2. Sia M una L -struttura infinita, allora, per ogni cardinalità infinita $\kappa \geq |L| + |M|$, c'è una estensione elementare N di M avente cardinalità $= \kappa$.

Dimostrazione. È facile usare la forma debole per salire sopra κ , e poi riscendere col teorema di Löwenheim-Skolem verso il basso. □

Esercizio 3.57 (Paradosso di Skolem). Se la teoria degli insiemi ZFC è coerente, allora ha un modello numerabile.

Esercizio 3.58 (Modelli numerabili di $\text{Th}(\mathbb{N}; 0, 1, +, \cdot, s)$). Dimostra che $\text{Th}(\mathbb{N}; 0, +, \cdot, s)$ ha 2^{\aleph_0} modelli numerabili (non isomorfi).

§3.8 Categoricità e completezza

Definizione 3.59 (Categoricità). Sia κ una cardinalità. Una L -teoria T si dice **κ -categorica** se esiste un unico modello di T di cardinalità κ , a meno di isomorfismi.

Proposizione 3.60 (Categorica infinita \implies completa)

Se una L -teoria T è κ -categorica con $|L| + \aleph_0 \leq \kappa$, allora T è completa.

Dimostrazione. Supponiamo il viceversa. Allora esistono φ e due modelli M_1, M_2 di T tali che $M_1 \models \varphi$ e $M_2 \models \neg\varphi$. Applicando i teoremi di Löwenheim-Skolem troviamo due L -strutture M'_1, M'_2 di cardinalità κ elementarmente equivalenti a M_1 e M_2 rispettivamente. Queste dovrebbero quindi essere due modelli di T che non possono essere isomorfi in quanto $M_1 \models \varphi$ e $M_2 \models \neg\varphi$. \square

Esercizio 3.61. La teoria T_{oldse} è completa.

Nota 3.62 — Abbiamo già dimostrato questo risultato per eliminazione dei quantificatori. Quello che segue è un argomento diretto.

Soluzione. Siccome c'è un solo ordine lineare denso e senza estremi numerabile (Teorema di isomorfismo di Cantor [INSERIRE RIFERIMENTO]), la teoria T_{oldse} è pertanto \aleph_0 -categorica, quindi, per la proposizione precedente, è completa. \square

Esercizio 3.63. Dai una assiomatizzazione esplicita delle seguenti teorie:

1. $\text{Th}(\mathbb{C}; 0, 1, +, \cdot)$;
2. $\text{Th}(\mathbb{Z}; s)$;
3. $\text{Th}(\mathbb{N}; s)$;
4. $\text{Th}(\mathbb{R}; 0, +)$.

Esercizio 3.64. Sia T la teoria nel linguaggio $L = \{\sim\}$ che dice che \sim è una relazione di equivalenza. Classifica le L -teorie complete che estendono T .

Soluzione. Ogni estensione completa di T è determinata dal numero di classi di equivalenza finite e dal numero di classi di equivalenza infinite. Infatti, siano M_1 e M_2 due modelli di T con lo stesso numero di classi di equivalenza finite e infinite, allora è facile costruire un isomorfismo tra i due modelli mappando le classi di equivalenza in modo opportuno. D'altra parte, se M_1 e M_2 hanno un diverso numero di classi di equivalenza finite o infinite, allora esiste una formula che distingue i due modelli (ad esempio, se M_1 ha n classi finite e M_2 ne ha $m \neq n$, la formula che dice "ci sono esattamente n classi finite" distingue i due modelli). \square

§4 Sintassi

Definizione 4.1 (Formula in ND). Sia L un linguaggio, una **L -formula in ND** è una sequenza finita di simboli nell'alfabeto $\{\wedge, \vee, \rightarrow, \neg, \forall, \exists, \top, \perp\} \cup L$ (il primo insieme è costituito dai **simboli logici**).

Definizione 4.2 (Dimostrazione in ND). Una **dimostrazione in ND** di una L -formula φ è una sequenza di L -formule $\varphi_1, \varphi_2, \dots, \varphi_n$ tale che $\varphi_n = \varphi$ e per ogni $1 \leq k \leq n$, o φ_k è un assioma, oppure φ_k è in T , oppure φ_k è ottenuta da formule precedenti tramite una **regola di inferenza** del **sistema di deduzione naturale** (ND) qui di seguito riportate.

§4.1 Sistema di deduzione naturale (ND)

Definizione 4.3 (Sequente in ND). Una **sequente** $T \vdash_{\text{ND}} \varphi$ indica che esiste una dimostrazione in ND di φ a partire dalle formule in T .

Definizione 4.4 (Formula in $\text{ND}_{\rightarrow, \perp, \exists, =}$). Sia L un linguaggio, una **L -formula in $\text{ND}_{\rightarrow, \perp, \exists, =}$** è una sequenza finita di simboli nell'alfabeto $\{\rightarrow, \perp, \exists, =\} \cup L$ (il primo insieme è costituito dai **simboli logici** di $\text{ND}_{\rightarrow, \perp, \exists, =}$).

Definizione 4.5 (Dimostrazione in $\text{ND}_{\rightarrow, \perp, \exists, =}$). Una **dimostrazione in $\text{ND}_{\rightarrow, \perp, \exists, =}$** di una L -formula φ è una sequenza di L -formule $\varphi_1, \varphi_2, \dots, \varphi_n$ tale che $\varphi_n = \varphi$ e per ogni $1 \leq k \leq n$, o φ_k è un assioma, oppure φ_k è in T , oppure φ_k è ottenuta da formule precedenti tramite una **regola di inferenza** del **sistema di deduzione naturale ridotto** ($\text{ND}_{\rightarrow, \perp, \exists, =}$) qui di seguito riportate.

§4.2 Sistema ridotto ($\text{ND}_{\rightarrow, \perp, \exists, =}$)

Definizione 4.6 (Sequente in $\text{ND}_{\rightarrow, \perp, \exists, =}$). Un **sequente** $T \vdash_{\text{ND}_{\rightarrow, \perp, \exists, =}} \varphi$ indica che esiste una dimostrazione in $\text{ND}_{\rightarrow, \perp, \exists, =}$ di φ a partire dalle formule in T .

Definizione 4.7 ($\neg, \top, \wedge, \vee, \forall$). Definiamo i seguenti simboli logici:

$$\begin{aligned}\neg \varphi &\stackrel{\text{def}}{=} \varphi \rightarrow \perp & \top &\stackrel{\text{def}}{=} \neg \perp & \varphi \wedge \psi &\stackrel{\text{def}}{=} \neg(\varphi \rightarrow \neg \psi) \\ \varphi \vee \psi &\stackrel{\text{def}}{=} \neg \varphi \rightarrow \psi & \forall x_k \varphi &\stackrel{\text{def}}{=} \neg \exists x_k \neg \varphi\end{aligned}$$

Esercizio 4.8 ($\text{ND}_{\rightarrow, \perp, \exists, =}$ dimostra ND). In $\text{ND}_{\rightarrow, \perp, \exists, =}$ si possono dimostrare le regole mancanti di ND nel seguente ordine: In_\neg , El_\perp , El_\neg , In_\top , In_\wedge , El_\wedge , In_\vee , El_\vee , In_\forall , El_\forall .

Esercizio 4.9 (Semantica di Tarski indotta). Verificare che la semantica di Tarski, già definita in precedenza, coincide con quella indotta dalle definizioni date sopra per i simboli: $\neg, \top, \wedge, \vee, \forall$.

Notazione 4.10 (Sequenti) — Avendo verificato che ND e $\text{ND}_{\rightarrow, \perp, \exists, =}$ sono equivalenti, d'ora in avanti indicheremo semplicemente con $T \vdash \varphi$ il sequente dimostrabile in entrambi i sistemi.

§4.3 Teoremi di correttezza e completezza

Proposizione 4.11

Data T una L -teoria e φ una L -formula, vale che $T \vdash \varphi \iff$ per ogni L -struttura $M = (D; i)$ e per ogni $v : \text{Var} \rightarrow D$ valutazione delle variabili, si ha $M \models \{v\}T \implies M \models \{v\}\varphi$.

Corollario 4.12 (Teorema di correttezza)

Sia T una L -teoria fatta da formule chiuse, e sia φ una L -formula, allora:

$$T \vdash \varphi \implies T \models \varphi$$

Corollario 4.13 (Teorema di completezza)

Sia T una L -teoria fatta da formule chiuse, e sia φ una L -formula, allora:

$$T \models \varphi \implies T \vdash \varphi$$

Riferimenti bibliografici

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

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