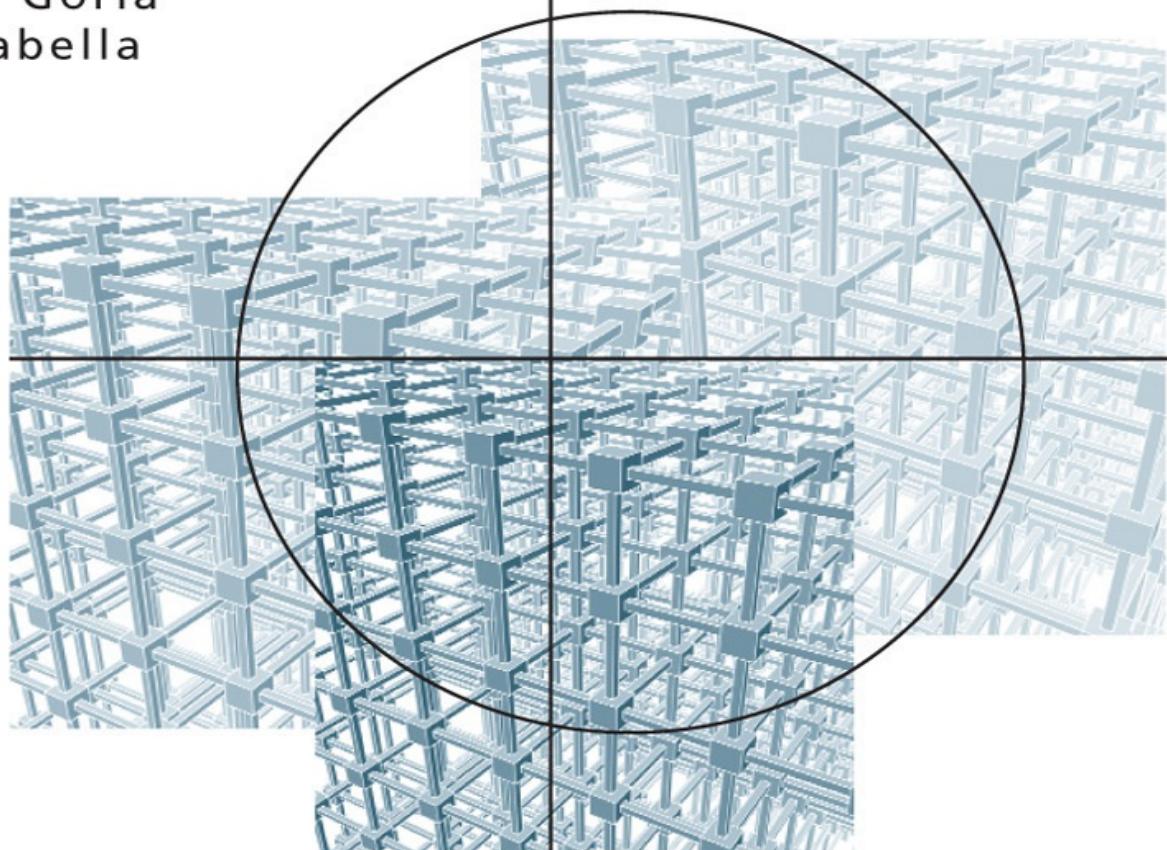


Giorgio T. Bagni
Daniele Gorla
Anna Labella



Introduzione alla logica e al linguaggio matematico

McGraw-Hill

Introduzione alla logica e al linguaggio matematico

Lo scopo di questo breve volume è guidare lo studente che si accinge a intraprendere un curriculum di studi scientifico nell'acquisire il linguaggio e il rigore di ragionamento matematico, indispensabili a chi voglia occuparsi e dedicarsi a una scienza.

Il volume quindi non vuole presentarsi come un manuale ma come un invito a capire perché le cose debbano essere fatte in un certo modo se si vuole che abbiano un certo valore scientifico. Nella trattazione si evitano quindi definizioni complesse e alcune dimostrazioni di teoremi preferendo

spiegare il motivo e la ragionevolezza di certe scelte, ricorrendo spesso a controesempi e, talvolta, a considerazioni di “buon senso”.

Altro grande pregio della presente opera è l'enorme corredo di esercizi, svolti e non svolti, di diversi gradi di difficoltà. Essi costituiscono uno strumento indispensabile per lo studente che, attraverso tentativi ed errori, voglia appropriarsi del linguaggio e del metodo matematico.

Sul sito web www.ateneonline.it/bagni sono disponibili le soluzioni di tutti gli *Esercizi da svolgere*.

collana di istruzione scientifica
serie di matematica

Giorgio T. Bagni
Daniele Gorla
Anna Labella

Introduzione alla logica e al linguaggio matematico

McGraw-Hill

Milano • New York • San Francisco • Washington D.C. • Auckland
Bogot • Lisboa • London • Madrid • Mexico City • Montreal
New Delhi • San Juan • Singapore • Sydney • Tokyo • Toronto

Copyright © 2014, McGraw-Hill
Education (Italy) S.r.l. Via Ripamonti, 89 -
20141 Milano.

Copyright edizione a stampa
©2010

The McGraw-Hill Companies, srl
Publishing Group Italia
Via Ripamonti, 89 - 20139 Milano



I diritti di traduzione, di riproduzione, di memorizzazione elettronica e di adattamento totale e parziale con qualsiasi mezzo (compresi i microfilm e le copie fotostatiche) sono riservati per tutti i Paesi.

La stampa dell'opera è consentita esclusivamente per uso personale.

Date le caratteristiche intrinseche di Internet, l'Editore non è responsabile per eventuali variazioni negli indirizzi e nei contenuti dei siti Internet riportati.

Nomi e marchi citati nel testo sono generalmente depositati o registrati dalle rispettive case produttrici.

Editor: Paolo Roncoroni

Produzione: Donatella Giuliani

Grafica di copertina: Editta Gelsomini

Realizzazione ePub: codeMantra

ISBN: 978-88-386-9175-1

*A Giorgio,
prematuramente scomparso
poco dopo la stesura di questo volume*

Indice

Prefazione

Parte A Introduzione alla matematica di base

1 Insiemi e operazioni su di essi

1.1 La nozione di “insieme”

1.2 Sottoinsiemi e inclusione

1.3 Operazioni sugli insiemi: unione,
intersezione, differenza e
complemento

1.4 Il prodotto cartesiano di due insiemi
Esercizi svolti

Esercizi da svolgere

2 Relazioni e loro proprietà

2.1 Sottoinsiemi del prodotto cartesiano

2.2 Relazioni tra un insieme e se stesso e
loro proprietà

2.3 Relazioni di equivalenza e insieme
quoziente

2.4 Relazioni d'ordine

Esercizi svolti

Esercizi da svolgere

3 Funzioni

3.1 La definizione di funzione

3.2 Funzioni iniettive, suriettive,
biettive

3.3 Composizione di funzioni

3.4 Funzioni parziali

3.5 La funzione identità e la funzione
inversa

Esercizi svolti

Esercizi da svolgere

4 Cardinalità

- 4.1 Equipotenza tra insiemi
 - 4.2 La potenza del numerabile
 - 4.3 La potenza del continuo
 - 4.4 Cenni sulle antinomie della teoria degli insiemi
 - 4.4.1 Le antinomie
 - 4.4.2 Gottlob Frege e Bertrand Russell
- Esercizi svolti
- Esercizi da svolgere

5 Numeri naturali

- 5.1 L'insieme dei numeri naturali
- 5.2 L'induzione
 - 5.2.1 Dimostrazioni e definizioni per induzione
 - 5.2.2 Varianti del Principio di Induzione: Induzione

strutturale e Induzione completa

5.3 Divisibilità e numeri primi

5.3.1 L'algoritmo euclideo e il massimo comun divisore

5.3.2 La rappresentazione dei numeri naturali

5.3.3 I numeri primi

Esercizi svolti

Esercizi da svolgere

6 Algebre di Boole

6.1 Proprietà di operazioni aritmetiche e insiemistiche

6.2 Algebre di Boole

Esercizi svolti

Esercizi da svolgere

Parte B Introduzione alla logica

matematica

7 La formalizzazione matematica

8 Calcolo degli enunciati

8.1 Enunciati, connettivi e valori di verità

8.1.1 Verità

8.1.2 Enunciati

8.1.3 Connnettivi e valori di verità

8.1.4 Interpretazioni, equivalenza logica e validità

8.1.5 Principii logici e ragionamento per assurdo

8.2 Il metodo dei tableau proposizionali

8.2.1 La confutazione di un enunciato composto

8.2.2 La costruzione di un tableau proposizionale

8.2.3 Correttezza e completezza

8.3 Il sistema deduttivo di Gentzen

- 8.3.1 Il sistema G
 - 8.3.2 Deduzione in G e tableau
- 8.4 Il sistema deduttivo di Hilbert
 - 8.4.1 Assiomi e Modus Ponens
 - 8.4.2 Regole derivate del sistema di Hilbert
- Esercizi svolti
- Esercizi da svolgere

9 Calcolo dei predicati

- 9.1 Logica dei predicati: calcolo dei quantificatori
 - 9.1.1 Dall'alfabeto alle formule predicative
 - 9.1.2 I quantificatori
 - 9.1.3 Variabili vincolate, variabili libere e sostituzioni
 - 9.1.4 Modelli e validità
- 9.2 Il metodo dei tableau per il calcolo dei predicati
 - 9.2.1 Tableau e quantificatori

9.2.2 Regole per la costruzione di un tableau predicativo

9.2.3 Correttezza e completezza del metodo dei tableau

9.3 Il sistema di Gentzen per il calcolo dei predicati

9.4 Cenni sul sistema di Hilbert per il calcolo dei predicati

9.5 Teorie

Esercizi svolti

Esercizi da svolgere

10 Procedimenti di risoluzione

10.1 La risoluzione per il calcolo degli enunciati

10.1.1 Forme clausali

10.1.2 Risoluzione

10.2 La risoluzione per il calcolo dei predicati

10.2.1 Forma prenessa e forma clausale

10.2.2 Unificazione
10.2.3 Risoluzione
10.2.4 Dimostrazione automatica
10.2.5 Programmazione logica
Esercizi svolti
Esercizi da svolgere

11 Prospettive e approfondimenti

Bibliografia

Indice analitico

Prefazione

Lo scopo di questo breve volume è ben lungi dall’essere modesto: esso vuole guidare uno studente che si accinge ad intraprendere un curriculum di studi scientifico nell’acquisire il linguaggio ed il rigore di ragionamento matematico che riteniamo indispensabili a chi voglia occuparsi e dedicarsi ad una scienza nel senso attuale del termine. Ora, tutti coloro che hanno esperienza in questo campo sanno quanto sia difficile trasmettere non tanto un contenuto, quanto una metodologia e, soprattutto, un modo di esprimersi e di pensare. Il giovane si aspetta di dover apprendere dei contenuti,

mentre noi dobbiamo condurlo a guardare certi contenuti con occhi ed atteggiamento completamente diversi da quelli cui è abituato nell'esperienza ordinaria. I contenuti sono, in parte, soltanto un oggetto sul quale esercitare questo nuovo atteggiamento.

Qualunque ramo dell'attività umana ha raggiunto ormai un grado di specializzazione tale da usare un linguaggio così specifico da non essere comprensibile a chi non vi è stato addestrato (si pensi ad esempio al linguaggio sportivo) e questo ci può aiutare nel far capire ai ragazzi quale sia il tipo di sforzo che si vuole da loro. Tuttavia il linguaggio matematico ha una tale complessità ed articolazione da richiedere obiettivamente un impegno non banale nella sua acquisizione.

La necessità poi di rispettare i tempi

dovuti all'organizzazione didattica degli insegnamenti non facilita di certo il maturare di questo cambiamento di prospettiva che, di fatto, ha impiegato tanti secoli per arrivare alla sofisticata forma che conosciamo oggi. Per dare appunto un'idea dell'evolversi del rigore matematico e non farlo sembrare come una legge assoluta ed immutabile nella sua astrazione, sono stati introdotti qua e là spunti storici che dovrebbero dare l'intuizione di un lento e laborioso evolversi di certi concetti, che porta con sé tutta la fatica ed il gusto della scoperta (o dell'invenzione, a seconda dei punti di vista).

Il volume quindi non vuole presentarsi come un manuale (esistono già numerose e pregevoli opere in tal senso a tutti i livelli), ma come un invito a capire perché le cose debbano essere fatte in un certo

modo se si vuole che abbiano un certo valore scientifico. Come si vedrà, sono state evitate definizioni complesse e alcune dimostrazioni di teoremi, mentre si è preferito “perdere tempo” nello spiegare il motivo e la ragionevolezza di certe scelte, ricorrendo spesso a controesempi e, talvolta, a considerazioni di “buon senso”. Dovendo fissare dei contenuti, si è scelto di prendere in esame la teoria ingenua degli insiemi, la struttura dei numeri naturali e gli elementi di base della logica matematica. Questo ci ha permesso di esaminare sul campo le diverse tecniche di dimostrazione comuni in matematica ed i principii ad esse soggiacenti, ma anche di fornire quelle basi matematiche che serviranno per insegnamenti più avanzati.

Avendo fatto la scelta di essere rigorosi ma non esaurienti nei diversi argomenti,

ci siamo appoggiati, per quanto riguarda la parte di logica, al testo di M. Ben-Ari “Logica matematica per informatica” [6], che costituisce, a nostro giudizio, un ottimo esempio di manuale didatticamente valido. Ad esso faremo frequenti riferimenti quando ometteremo qualche dimostrazione; ad esso indirizziamo gli studenti più volenterosi, speriamo stimolati da questa nostra introduzione, per ampliare le loro conoscenze, anche se ancora ad un livello elementare. A tale scopo, riportiamo in bibliografia anche altre interessanti letture che caldamente suggeriamo come approfondimento.

Il presente volume è stato scritto per studenti del primo anno di Informatica, dopo una pluriennale esperienza in questo campo da parte di tutti gli autori, manon è esclusivamente dedicato ad essi. Pensiamo

che possa essere utile, come dicevamo, a tutti coloro che affrontano un curriculum universitario di carattere scientifico.

Riteniamo che uno dei principali pregi della presente opera sia l'enorme corredo di esercizi, svolti e non svolti, e di diversi gradi di difficoltà. Essi costituiscono, a nostro parere, uno strumento indispensabile per lo studente che, attraverso tentativi ed errori, voglia appropriarsi del linguaggio e del metodo matematico. Un piccolo capitolo del libro è stato infatti destinato alla formalizzazione, che è, in fondo, centrale in questo nostro discorso; tuttavia, riteniamo che soltanto la pratica possa veramente far conseguire allo studente questa capacità.

Ringraziamo Francesco Davi per il meticoloso lavoro di rilettura e segnalazione di errori presenti in una

versione preliminare del testo.

Gli autori

Roma, Giugno 2009

Ringraziamenti dell'Editore

L'Editore ringrazia il professor Pier Luigi Ferrari dell'Università degli Studi del Piemonte Orientale, che ha partecipato alla review del testo e che con le sue preziose indicazioni ha contribuito alla realizzazione di *Introduzione alla logica e al linguaggio matematico*.

Parte A

Introduzione alla matematica di base

Insiemi e operazioni su di essi

1.1 La nozione di “insieme”

Dal punto di vista intuitivo, il concetto di insieme può essere fatto corrispondere all’atto mentale mediante il quale associamo alcuni elementi in un tutto unico detto *insieme* (o *classe*). Tale descrizione può essere considerata vaga e imprecisa; ma in vista degli scopi di questo volume, possiamo attenerci a questo punto di vista intuitivo per introdurre le principali nozioni insiemistiche.¹ Non è richiesta alcuna

particolare omogeneità tra gli elementi che costituiscono un insieme: è possibile associare nello stesso insieme un numero qualsiasi di elementi di qualsiasi genere. **La definizione dell'insieme potrà avvenire attraverso l'elencazione dei suoi elementi o specificando una proprietà della quale i medesimi devono godere.**

Esempio 1.1. È possibile parlare di un insieme cui appartengono il nome del monte più alto della Terra, l'ultima lettera dell'alfabeto italiano e le soluzioni dell'equazione: $x^2 + 6 = 5x$. Tale insieme ha elementi: *Everest*, Z, 2 e 3. È anche possibile parlare dell'insieme dei “numeri reali maggiori di 4”.

Come si vede, il primo modo, più diretto, non sarebbe in grado di descrivere il secondo insieme, perché non sapremmo elencare i suoi elementi, mentre il secondo

richiede che la proprietà (il predicato “essere maggiore di”) sia formulata su un insieme preesistente; in questo caso quello dei numeri reali. Comunque sia definita, affinché una collezione di elementi possa essere considerata come un insieme vero e proprio, deve sempre essere possibile stabilire se un qualsiasi elemento appartiene (o non appartiene) a essa.

Esempio 1.2. *Ha senso, nella teoria degli insiemi, parlare dell’insieme costituito dai numeri reali maggiori di 4. Infatti è possibile stabilire oggettivamente se un qualsiasi elemento appartiene o no a tale insieme: per appartenere all’insieme, un elemento deve (contemporaneamente) essere un numero reale ed essere maggiore di 4. Dunque all’insieme introdotto apparterranno certamente elementi come 59 , 19 , $\frac{33}{8}$; mentre non vi apparterranno elementi*

come -1 , $\frac{31}{8}$, 4 (che so-no numeri reali, ma non sono maggiori di 4), o come Firenze, Hemingway, un esagono regolare, il numero $\sqrt{-1}$ (che non è un numero reale).

(Contro)esempio. *Non ha senso, nella teoria degli insiemi, parlare dell'insieme costituito dai libri interessanti: non è infatti possibile affermare oggettivamente se un libro è interessante oppure se non lo è.*

È convenzione spesso accettata indicare gli insiemi con lettere maiuscole (A, B, C, \dots) e gli elementi con minuscole (a, b, c, \dots). L'appartenenza dell'elemento a all'insieme A si indica con la scrittura $a \in A$ nella quale il simbolo “ \in ” si legge: “appartiene”. La non appartenenza di b a B si indica con $b \notin B$. Quindi la condizione richiesta per poter parlare di insieme, espressa precedentemente, è: affinché I

sia un insieme, è richiesto che, per ogni elemento a , sia possibile stabilire che l'affermazione $a \in I$ sia *vera* o *falsa*.

Per indicare gli insiemi, con i loro elementi, si possono scegliere, in corrispondenza ai due modi di definirli, due procedimenti. Un primo metodo, detto *rappresentazione tabulare*, consiste nell'elencare tutti gli elementi che costituiscono l'insieme in questione tra parentesi graffe; per esempio, la scrittura $I = \{1, 2, 5\}$ significa che all'insieme I appartengono (solamente) gli elementi 1, 2 e 5. Notiamo che l'ordine con cui si elencano gli elementi è privo di importanza. Un insieme non presuppone alcun particolare ordinamento dei suoi elementi e nemmeno considera eventuali ripetizioni fatte nell'elencarli; le scritture

$$\{1, 2, 5\} \quad \{1, 5, 2\} \quad \{2, 1, 5\} \quad \{2, 5, 1\} \quad \{5, 1, 2\}$$

$$\{5, 2, 1\} \quad \{1, 2, 1, 5\}$$

identificano tutte lo stesso insieme, cioè l'insieme avente per elementi 1, 2 e 5.

Talvolta la rappresentazione tabulare può essere scomoda e addirittura impraticabile quando l'insieme in questione è costituito da infiniti elementi. La *rappresentazione caratteristica* si ottiene evidenziando (ove ciò sia possibile) una condizione necessaria e sufficiente per l'appartenenza di un elemento all'insieme considerato. La scrittura dell'insieme avrà forma:

$$\{x : x \text{ rispetta un'assegnata condizione}\}$$

nella quale il simbolo “:” (talvolta sostituito da “|”) si legge “tale che”.²

Esempio 1.3. Indichiamo l'insieme I di

interi il cui quadrato è minore di 15:

rappresentazione tabulare: $I = \{-3, -2, -1, 0, 1, 2, 3\}$;

rappresentazione caratteristica: $I = \{x: x \text{ è un numero intero e } -3 \leq x \leq 3\}$.

Si noti, inoltre, che si possono trovare diverse (ma equivalenti) rappresentazioni caratteristiche per lo stesso insieme.

Nell'esempio precedente, avremmo alternativamente potuto caratterizzare l'insieme scrivendo $I = \{x: x \text{ è un numero intero e } x^2 \leq 15\}$. Sempre nell'esempio precedente abbiamo scritto (usando parole tratte dalla lingua italiana) che l'elemento x “è un numero intero”, cioè appartiene all'*insieme costituito dai numeri interi*. Questo insieme è generalmente indicato con il simbolo \mathbb{Z} . Anche altri insiemi numerici sono di uso frequente:

| | |
|----------------|---|
| \mathbb{N} | insieme dei numeri naturali; |
| \mathbb{N}_0 | insieme dei numeri naturali non nulli; |
| \mathbb{Z} | insieme dei numeri interi; |
| \mathbb{Z}_0 | insieme dei numeri interi non nulli; |
| \mathbb{Z}_+ | insieme di numeri interi positivi; |
| \mathbb{Z}_- | insieme dei numeri interi negativi; |
| \mathbb{Q} | insieme dei numeri razionali; |
| \mathbb{Q}_0 | insieme dei numeri razionali non nulli; |
| \mathbb{Q}_+ | insieme dei numeri razionali positivi; |
| \mathbb{Q}_- | insieme dei numeri razionali negativi; |
| \mathbb{R} | insieme dei numeri reali; |
| \mathbb{R}_0 | insieme dei numeri reali non nulli; |
| \mathbb{R}_+ | insieme dei numeri reali positivi; |
| \mathbb{R}_- | insieme dei numeri reali negativi; |
| \mathbb{C} | insieme dei numeri complessi. |

Dunque l'insieme dei numeri interi il cui quadrato è minore di 15 può scriversi $\{x: x \in \mathbb{Z} \text{ e } -3 \leq x \leq 3\}$ oppure $\{x \in \mathbb{Z}: -3 \leq x \leq 3\}$.

Come si vede, la condizione con la quale si definisce un insieme è di solito espressa mediante una (o più) proprietà o “predicato” da soddisfare. Questo fatto è il legame tra la teoria degli insiemi e la logica degli enunciati e dei prediciati che vedremo più avanti. L'insieme privo di

elementi si dice *insieme vuoto*; si indica col simbolo \emptyset .

Esempio 1.4. *L'insieme costituito dalle soluzioni intere di $x^3 = 2$ è l'insieme vuoto, \emptyset : ciò equivale ad affermare che l'equazione data non ha radici intere.*

Attenzione! Due insiemi possono essere considerati uguali anche se la loro definizione, data mediante una proprietà, è differente. Si pensi, per esempio all'*insieme dei cerchi che sono anche quadrati* e l'*insieme degli uomini viventi che hanno compiuto 500 anni*.

È chiaro che, per motivi diversi, questi due insiemi hanno gli stessi elementi: in questo caso sono entrambi vuoti e saranno per noi lo stesso insieme. Infatti consideriamo *uguali* due insiemi se hanno gli stessi elementi e scriveremo $A = B$. Questa (apparentemente) innocente

definizione ha delle gravi conseguenze: accettandola, consideriamo un insieme soltanto in base ai suoi elementi (estensione) e non, per esempio, alla proprietà che lo definisce (intensione).

1.2 Sottoinsiemi e inclusione

Dato un insieme A , diremo che un insieme B è un *sottoinsieme* di A se a B non appartengono elementi non appartenenti ad A : dunque B può essere costituito da alcuni elementi di A , o da tutti gli elementi di A , o da nessun elemento.

Definizione 1.1. *L'insieme B si dice sottoinsieme dell'insieme A se ogni elemento di B è elemento di A , cioè se (un generico) $x \in B$ allora $x \in A$; si dice anche che B è incluso in A e si scrive $B \subseteq$*

A.

Tra tutti i sottoinsiemi di un insieme dato A troviamo sempre l'insieme A stesso e l'insieme vuoto \emptyset : essi sono detti *sottoinsiemi impropri* di A ; un sottoinsieme di A diverso da A stesso e da \emptyset si dice *sottoinsieme proprio* di A .

L'insieme vuoto ammette uno e un solo sottoinsieme (improprio): \emptyset . Un insieme costituito da un solo elemento, $A = \{a\}$, ammette due sottoinsiemi impropri, \emptyset e A stesso, e non ammette alcun sottoinsieme proprio.

Definizione 1.2. Si dice insieme delle parti di un insieme A l'insieme $\wp(A)$ avente per elementi tutti i sottoinsiemi (propri e impropri) di A : $\wp(A) = \{B: B \subseteq A\}$.

Per quanto detto, quando A è sottoinsieme di B e contemporaneamente

B è sottoinsieme di A (cioè quando $A \subseteq B$ e $B \subseteq A$), allora i due insiemi A e B sono uguali.

1.3 Operazioni sugli insiemi: unione, intersezione, differenza e complemento

Definizione 1.3. L'insieme $A \cup B$, unione degli insiemi A e B , è l'insieme al quale appartengono gli elementi che appartengono almeno a uno degli insiemi A e B : $A \cup B = \{x : x \in A \text{ o } x \in B\}$.

Esempio 1.5. L'unione degli insiemi $A = \{x \in \mathbb{R} : 1 < x < 5\}$ e $B = \{x \in \mathbb{R} : 3 < x < 15\}$ è l'insieme $A \cup B = \{x \in \mathbb{R} : 1 < x < 15\}$.

Definizione 1.4. L'insieme $A \cap B$, intersezione degli insiemi A e B , è

l'insieme al quale appartengono gli elementi che appartengono contemporaneamente agli insiemi A e B :
$$A \cap B = \{x : x \in A \text{ e } x \in B\}.$$

Esempio 1.6. *L'intersezione di $A = \{x \in \mathbb{R} : 1 < x < 5\}$ e $B = \{x \in \mathbb{R} : 3 < x < 15\}$ è l'insieme $A \cap B = \{x \in \mathbb{R} : 3 < x < 5\}$.*

Due insiemi a intersezione vuota si dicono *disgiunti*. Si verifica che:

- $I \cup I = I$;
- $I \cup \emptyset = I$;
- $I \cap I = I$;
- $I \cap \emptyset = \emptyset$;
- $I \cup J = J \cup I$ (proprietà commutativa);
- $I \cap J = J \cap I$ (proprietà commutativa);
- $I \cup (J \cup K) = (J \cup I) \cup K$ (proprietà associativa);
- $I \cap (J \cap K) = (J \cap I) \cap K$ (proprietà

associativa);

- $I \cup (J \cap K) = (I \cup J) \cap (I \cup K)$

(proprietà distributiva);

- $I \cap (J \cup K) = (I \cap J) \cup (I \cap K)$

(proprietà distributiva).

Definizione 1.5. L'insieme $A \setminus B$, differenza degli insiemi A e B , è l'insieme al quale appartengono gli elementi che appartengono ad A ma non appartengono a B : $A \setminus B = \{x : x \in A \text{ e } x \notin B\}$.

Esempio 1.7. La differenza di $A = \{x \in \mathbb{R} : 1 < x < 5\}$ e $B = \{x \in \mathbb{R} : 3 < x < 15\}$ è l'insieme $A \setminus B = \{x \in \mathbb{R} : 1 < x \leq 3\}$.

Per quanto riguarda le proprietà dell'operazione introdotta, si verifica che:

- $I \setminus I = \emptyset$;

- $I \setminus \emptyset = I$;

- $\emptyset \setminus I = \emptyset$.

Si può pensare di fissare una volta per tutte (almeno all'interno di un certo discorso) l'insieme I e chiamarlo *universo*, allora avremo la seguente:

Definizione 1.6. *Il complemento di un sottoinsieme A di I è il sottoinsieme $\overline{A} = \{x \in I : x \notin A\}$.*

Dalla definizione si possono ricavare per esercizio le seguenti proprietà:

- $\overline{\overline{A}} \cap A = \emptyset$;
- $\overline{\overline{A}} \cup A = I$;
- $\overline{\overline{A}} = A$.

È da notare che le operazioni tra insiemi che abbiamo introdotto sono state espresse in base al predicato di appartenenza (\in) e usando i connettivi

grammaticali *o*, *e* e *non*, combinando cioè tra di loro mediante connettivi, le proprietà che definiscono gli insiemi. In realtà anche l'inclusione tra insiemi è stata introdotta attraverso il *se... allora*.

1.4 Il prodotto cartesiano di due insiemi

Quanto finora esposto a proposito del concetto di insieme non fa riferimento all'ordine con cui gli elementi di un insieme sono elencati. È però utile, in determinati casi, specificare un particolare ordinamento all'interno di un dato insieme: è quanto ci accingiamo a fare introducendo la *coppia ordinata*.

Siano A e B due insiemi che supporremo dati. Una coppia ordinata di elementi di A e di B si indica con il simbolo (a, b) . Intuitivamente, essa è un particolare

insieme costituito da due elementi, il *primo* dei quali apparteneente ad A , il *secondo* a B . Più precisamente, la coppia ordinata (a, b) può essere introdotta sulla base del concetto di coppia ordinaria (cioè non ordinata) $\{a, b\}$: si può per esempio scegliere due elementi distinti esterni all'insieme considerato (nel nostro caso, agli insiemi A e B), per esempio 1 e 2, e definire (a, b) come $\{\{1, a\}, \{2, b\}\}$.
Alternativamente, si può “specificare” quale dei due elementi è da ritenere il primo della coppia ordinata, considerando per (a, b) per esempio $\{\{a\}, \{a, b\}\}$.

Definizione 1.7. *L'insieme $A \times B$, prodotto cartesiano degli insiemi A e B , è l'insieme avente per elementi tutte le coppie ordinate (a, b) , con $a \in A$ e $b \in B$:*
$$A \times B = \{(a, b): a \in A \text{ e } b \in B\}.$$

Nel caso in cui (almeno) uno dei due insiemi A e B sia \emptyset , $A \times B$ è vuoto.

Esempio 1.8. Il prodotto cartesiano di $A = \{c, d\}$ e $B = \{2, 4, 7\}$ è $\{(c, 2), (d, 2), (c, 4), (d, 4), (c, 7), (d, 7)\}$.

La costruzione del prodotto cartesiano di due insiemi A e B è importante perché combina gli elementi dei due insiemi in modo tale da poter riottenere per ogni elemento dell'uno tutta una copia dell'altro. Introduciamo dunque le *proiezioni*.

Definizione 1.8. Sia $S \subseteq A \times B$. Si dice proiezione di S su A (rispettivamente, su B) l'insieme $\{x: x \in A \text{ e } (x, b) \in S, \text{ per almeno un } b \in B\}$ (rispettivamente, $\{y: y \in B \text{ e } (a, y) \in S, \text{ per almeno un } a \in A\}$).

Evidentemente la proiezione di tutto

l'insieme (non vuoto) $A \times B$ su A è A stesso, e su B è B stesso. Possiamo restringere le proiezioni a un sottoinsieme $S \subseteq A \times B$ (l'importanza dei sottoinsiemi del prodotto cartesiano apparirà evidente a partire dal prossimo capitolo).

Esempio 1.9. Sia S il sottoinsieme di $\mathbb{N} \times \mathbb{N}$ costituito dalle coppie (n, m) tali che $n+m = 3$. Lasciamo al lettore di verificare che $S = \{(0, 3), (1, 2), (2, 1), (3, 0)\}$. La proiezione di S su entrambe le sue componenti è $\{0, 1, 2, 3\}$.

Nell'esempio precedente, \mathbb{N} è sia il primo che il secondo degli insiemi dei quali viene considerato il prodotto cartesiano $\mathbb{N} \times \mathbb{N}$. Tuttavia, in generale, tali insiemi saranno distinti, come appare chiaro dall'esempio seguente.

Esempio 1.10. Sia T il sottoinsieme di \mathbb{N}

$\times \mathbb{R}$ costituito dalle coppie (n, r) tali che $r = \sqrt{n}$. La proiezione di T sulla prima componente è \mathbb{N} mentre la proiezione di T sulla seconda componente è $\{r \in \mathbb{R}_+: r^2 \in \mathbb{N}\}$.

Avendo introdotto un ordine nelle coppie, il prodotto $A \times B$ sarà diverso da $B \times A$, perché i suoi elementi saranno diversi; ciò non toglie che fra i due insiemi ci sia una certa somiglianza (che vedremo in seguito), che sfruttiamo per esempio quando consideriamo indifferentemente una lista di persone indicata per nome e cognome oppure per cognome e nome. Tuttavia occorre sapere quando stiamo usando una lista o l'altra per evitare ambiguità.

La considerazione di coppie può essere estesa a quella di terne, quadruple o, in generale, n -ple, iterando la costruzione del

prodotto cartesiano, costruendo cioè l'insieme $(\dots(A_1 \times A_2) \times \dots) \times A_n$.

Esercizi svolti

Esercizio 1.1. *Dato $M = \{1, 2, 3\}$, determinare se le seguenti proposizioni sono corrette e perché.*

$$(a) 1 \in M$$

$$(b) 1 \subset M$$

$$(c) \{1\} \in M$$

$$(d) \{1\} \subset M$$

Soluzione: La (a) è corretta: 1 è un elemento di M e quindi 1 appartiene a M ,

mentre l'insieme che contiene 1 non è un elemento di M (e quindi la (c) è sbagliata). La (b) è sbagliata: essendo un elemento di M , 1 non ne è un sottoinsieme; invece $\{1\}$ è un sottoinsieme di M , e quindi la (d) è corretta.

Esercizio 1.2. *Quali dei seguenti insiemi sono uguali?*

1. *$\{x: x$ è una lettera della parola “reattore”};*
2. *l’insieme delle lettere che compaiono nella parola “teatro”};*
3. *$\{x: x$ è una lettera della parola “attore”};*
4. *l’insieme formato dalle lettere a, e, o, r, t.*

Soluzione: Due insiemi sono uguali se contengono gli stessi elementi. Pertanto, gli insiemi dati sono tutti uguali.

Ricordiamo che, quando lavoriamo con un insieme, né l'ordine in cui i suoi elementi sono elencati né eventuali ripetizioni di uno stesso elemento contano.

Esercizio 1.3. *Quali dei seguenti insiemi sono differenti: \emptyset , $\{0\}$, $\{\emptyset\}$?*

Soluzione: Sono tutti diversi: \emptyset è un insieme senza elementi; $\{0\}$ è un insieme che contiene un solo elemento, il numero 0; $\{\emptyset\}$ è anch'esso un insieme che contiene un solo elemento, ma questo elemento è a sua volta un insieme (in particolare, l'insieme vuoto).

Esercizio 1.4. *Siano $A = \{1, 2, 3, 4\}$, $B = \{2, 4, 6, 8\}$, e $C = \{3, 4, 5, 6\}$. Trovare:*

- (a) $A \cup B$
- (b) $B \cup C$
- (c) $(A \cup B) \cup C$

- (d) $A \cap B$
- (e) $B \cap B$
- (f) $(A \cap B) \cap C$
- (g) $A \setminus B$
- (h) $B \setminus B$
- (i) $(A \setminus B) \setminus C$

Soluzione: L'unione di due insiemi X e Y è l'insieme che contiene tutti gli elementi di X e tutti gli elementi di Y (senza ripetizioni e in un qualsiasi ordine); pertanto, $A \cup B = \{1, 2, 3, 4, 6, 8\}$, $B \cup B = \{2, 4, 6, 8\}$ e $(A \cup B) \cup C = \{1, 2, 3, 4, 5, 6, 8\}$.

L'intersezione di due insiemi X e Y è l'insieme che contiene tutti gli elementi comuni a X e Y ; pertanto, $A \cap B = \{2, 4\}$, $B \cap B = \{2, 4, 6, 8\}$ e $(A \cap B) \cap C = \{4\}$.

Infine, la differenza di due insiemi X e Y è l'insieme che contiene tutti gli elementi di X che non appartengono a Y ;

pertanto, $A \setminus B = \{1, 3\}$, $B \setminus A = \emptyset$ e $(A \setminus B) \setminus C = \{1\}$.

Esercizio 1.5. Sia $U = \{a, b, c, d, e\}$ un insieme universo e siano $A = \{a, b, d\}$ e $B = \{b, d, e\}$ due insiemi definiti in tale universo. Trovare:

- (a) \overline{B}
- (b) $\overline{A} \cap B$
- (c) $A \cup \overline{B}$
- (d) $\overline{A} \cap \overline{B}$
- (e) $\overline{B} \setminus \overline{A}$
- (f) $\overline{(A \cap B)}$

Soluzione: Il complementare di un insieme X rispetto a un universo U è definito come la differenza tra U e X , scritta $U \setminus X$; pertanto, $\overline{B} = \{a, c\}$, $\overline{A} \cap B = \{e\}$, $A \cup \overline{B} = \{a, b, c, d\}$, $\overline{A} \cap \overline{B} = \{c\}$, $\overline{B} \setminus \overline{A} = \{a\}$ e $\overline{(A \cap B)} = \{a, c, e\}$.

Esercizio 1.6. Dimostrare la proprietà commutativa dell'unione insiemistica,

cioè che, comunque si prendano due insiemi A e B , si ha che $A \cup B = B \cup A$.

Soluzione: Useremo qui e in seguito una tecnica di prova standard per dimostrare l'uguaglianza di due insiemi X e Y . Come già più volte detto, X e Y sono lo stesso insieme se hanno gli stessi elementi; pertanto, per dimostrare che X è uguale a Y , dobbiamo dimostrare che ogni elemento di X è anche un elemento di Y e viceversa (in simboli, che $X \subseteq Y$ e $Y \subseteq X$). Questa tecnica di prova è chiamata *doppia inclusione*.

Per definizione $A \cup B = \{x: x \in A \text{ o } x \in B\}$ e $B \cup A = \{x: x \in B \text{ o } x \in A\}$. Si deve dimostrare che $A \cup B \subseteq B \cup A$ e che $B \cup A \subseteq A \cup B$. Se $x \in A \cup B$, allora $x \in A$ oppure $x \in B$; quindi $x \in B \cup A$. L'inclusione opposta è simmetrica.

Esercizio 1.7. *Dimostrare che \emptyset è*

l'elemento neutro per l'unione, cioè che, comunque si prenda un insieme A , si ha che $A \cup \emptyset = A$.

Soluzione: Per definizione $A \cup \emptyset = \{x: x \in A \text{ o } x \in \emptyset\}$. Per definizione $x \in \emptyset$ non è mai verificata; quindi $\{x: x \in A \text{ o } x \in \emptyset\} = \{x: x \in A\} = A$, da cui $A \cup \emptyset = A$.

Esercizio 1.8. *Dimostrare che la differenza insiemistica non gode della proprietà commutativa.*

Soluzione: Se valesse, la proprietà commutativa della differenza insiemistica affermerebbe che, *comunque si prendono* due insiemi A e B , si ha che $A \setminus B = B \setminus A$. Per alcune scelte di A e B questo vale: per esempio, basta prendere $A = B$; ma in generale questo non vale. Si prendano per esempio $A = \{1, 2, 3\}$ e $B = \{3, 4\}$: per definizione, $A \setminus B = \{1, 2\}$ mentre $B \setminus A =$

$\{4\}$. Poiché abbiamo mostrato che la proprietà non è verificata per un'opportuna scelta di A e B , otteniamo che la differenza insiemistica non è commutativa.

Esercizio 1.9. Siano $A = \{1, (2, 3)\}$ e $B = \{2, 4\}$. Si calcoli il prodotto cartesiano $A \times B$.

Soluzione: Il prodotto cartesiano di due insiemi X e Y è formato da tutte e sole le coppie (x, y) tali che x è un qualsiasi elemento di X e y è un qualsiasi elemento di Y . Pertanto, $A \times B = \{(1, 2), (1, 4), ((2, 3), 2), ((2, 3), 4)\}$. Si noti, infatti, che A contiene *due* elementi: uno è il numero 1, l'altro è la coppia $(2, 3)$.

Esercizio 1.10. Si calcoli il prodotto cartesiano tra l'insieme $A = \{1, 2, 3\}$ e l'insieme $B = \{x: \frac{x}{0} = 3\}$.

Soluzione: Essendo l'insieme B vuoto, si ha che anche $A \times B$ è vuoto.

Esercizio 1.11. *Si dimostri che il prodotto cartesiano non è commutativo.*

Soluzione: Consideriamo $A = \{1\}$ e $B = \{2\}$; allora, $A \times B = \{(1, 2)\}$ e $B \times A = \{(2, 1)\}$. Si osservi che la coppia $(1, 2)$ è *diversa* dalla coppia $(2, 1)$: infatti, mentre negli insiemi l'ordine con cui gli elementi vengono dati non conta, nelle coppie (e, in generale, nelle n -ple) l'*ordine conta!* Quindi, $A \times B \neq B \times A$.

Esercizi da svolgere

Esercizio 1.12. *Sia $A = \{2, \{4, 5\}, 4\}$. Quali delle seguenti proposizioni sono sbagliate e perché?*

- (a) $\{4, 5\} \subset A$
- (b) $\{4, 5\} \in A$
- (c) $\{\{4, 5\}\} \subset A$
- (d) $5 \in A$
- (e) $\{5\} \in A$
- (f) $\{5\} \subset A$

Esercizio 1.13. Sia $B = \{1, 0\}$. Dire se ciascuna delle seguenti proposizioni è giusta o sbagliata:

- (a) $\{0\} \in B$
- (b) $\emptyset \in B$
- (c) $\{0\} \subset B$
- (d) $0 \in B$
- (e) $0 \subset B$

Esercizio 1.14. Siano $A = \{1, 2, 3, 4\}$, $B = \{2, 4, 6, 8\}$, e $C = \{3, 4, 5, 6\}$. Trovare:

- (a) $A \cup C$

- (b) $B \cup C$
- (c) $C \cup C$
- (d) $A \cup (B \cup C)$
- (e) $A \cap C$
- (f) $B \cap C$
- (g) $C \cap C$
- (h) $A \cap (B \cup C)$
- (i) $A \setminus C$
- (l) $B \setminus C$
- (m) $C \setminus C$
- (n) $A \setminus (B \setminus C)$

Esercizio 1.15. Sia $U = \{a, b, c, d, e, f, g\}$ un insieme universo e siano $A = \{a, b, c, d, e\}$, $B = \{a, c, e, g\}$ e $C = \{b, e, f, g\}$ tre insiemi definiti in tale universo. Trovare:

- (a) \overline{B}
- (b) $\overline{A} \setminus B$
- (c) $\overline{B} \cup C$

(d) $\overline{(A \setminus C)}$

(e) $\overline{C} \cap A$

(f) $\overline{(A \setminus \overline{B})}$

(g) $\overline{(A \cap \overline{A})}$

Esercizio 1.16. *Dimostrare per doppia inclusione la proprietà commutativa dell'intersezione insiemistica.*

Esercizio 1.17. *Dimostrare per doppia inclusione la proprietà associativa dell'intersezione insiemistica, cioè che, comunque si prendano tre insiemi A , B e C , si ha che $A \cap (B \cap C) = (A \cap B) \cap C$.*

Esercizio 1.18. *Dimostrare che \emptyset è l'elemento azzeratore dell'intersezione, cioè che, comunque si prenda un insieme A , si ha che $A \cap \emptyset = \emptyset$.*

Esercizio 1.19. *Dimostrare che la differenza insiemistica non gode della*

proprietà associativa.

Esercizio 1.20. *Dati gli insiemi $A = \{x \in \mathbb{Z} : -1 \leq x \leq 2\}$ e $B = \{x \in \mathbb{Z} : 3 < x < 6\}$, si calcolino:*

- (a) $A \times B$
- (b) $B \times A$
- (c) $A \times A$
- (d) $B \times B$

Esercizio 1.21. *Si dimostri (per doppia inclusione) che $A \times (B \cup C) = (A \times B) \cup (A \times C)$.*

Esercizio 1.22. *Si dimostri (per doppia inclusione) che $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.*

¹ La teoria degli insiemi può essere introdotta assiomaticamente: l'applicazione del metodo assiomatico alla teoria degli insiemi ebbe in Ernest Zermelo (1871-1953) il principale protagonista.

² Il fatto che $\{x: x \text{ rispetta un'assegnata condizione}\}$ costituisca senz'altro un insieme riflette il "Principio di Comprensione" che caratterizza l'impostazione insiemistica di Gottlob Frege (1848-1925). Come avremo occasione di notare (nel paragrafo 4.4), tale affermazione non è per nulla scontata.

Relazioni e loro proprietà

2.1 Sottoinsiemi del prodotto cartesiano

Il nome di prodotto cartesiano è già di per sé suggestivo: ci ricorda infatti il piano cartesiano nel quale i punti sono individuati dalle coppie di coordinate che variano liberamente entrambe sull'insieme dei numeri reali. Il piano cartesiano, però, viene spesso utilizzato per descrivere il *luogo* di quei punti che soddisfano una certa proprietà: una parabola di equazione $y = x^2$ è un

sottoinsieme del piano costituito da quei punti per i quali ordinata ed ascissa non sono più qualunque, ma sono legate dalla condizione imposta dall'equazione. Come già sappiamo, una proprietà individua un sottoinsieme, che, in questo caso, siamo in grado di rappresentare graficamente. Il legame tra le coordinate imposto dall'equazione è una relazione fra esse.

In generale, siano dati gli insiemi A , B e $A \times B$. Come sappiamo, ad $A \times B$ appartengono tutte le coppie ordinate costituite da un primo elemento tratto da A e da un secondo elemento tratto da B . Ogni sottoinsieme di $A \times B$ potrà essere considerato come una relazione tra gli elementi di A e quelli di B esplicitata magari attraverso una qualche proprietà o una legge.

Definizione 2.1. Si dice relazione (o corrispondenza) tra gli insiemi A e B un

sottoinsieme del prodotto cartesiano $A \times B$.

Esempio 2.1. Consideriamo gli insiemi $A = \{Arno, Po, Tevere\}$ e $B = \{Firenze, Pisa, Torino\}$, e il loro prodotto cartesiano:

$$A \times B = \{ (Arno, Firenze), (Arno, Pisa), (Arno, Torino), (Po, Firenze), (Po, Pisa), (Po, Torino), (Tevere, Firenze), (Tevere, Pisa), (Tevere, Torino) \}$$

Tra tutte le coppie aventi per primo elemento un elemento di A (in questo caso, un fiume) e per secondo un elemento di B (una città), individuiamo quelle costituite dal nome di un fiume e da quello di una città bagnata da tale fiume:

$$R = \{ (Arno, Firenze), (Arno, Pisa), (Po,$$

Torino)}

Il sottoinsieme R di $A \times B$ è caratterizzato esattamente dalla proprietà che tutte (e soltanto) le coppie ad esso appartenenti sono costituite da un fiume e da una città bagnata da esso. Il sottoinsieme R ora indicato è uno dei possibili sottoinsiemi di $A \times B$: altri sottoinsiemi possono essere individuati da altre condizioni (tutte ugualmente valide dal punto di vista insiemistico).

Quando si considera la coppia (a, b) appartenente ad un dato sottoinsieme R di $A \times B$, si dice che l'elemento $a \in A$ ha per corrispondente $b \in B$ nella relazione R , oppure semplicemente che a è in relazione con b .

Una relazione, come ogni sottoinsieme del prodotto cartesiano di insiemi, può

spesso essere rappresentata graficamente. Il precedente esempio può essere rappresentato dalla seguente tabella:

| | Arno | Po | Tevere |
|---------|------|----|--------|
| Torino | | ✓ | |
| Pisa | ✓ | | |
| Firenze | ✓ | | |

Data una relazione tra gli elementi dell'insieme A e quelli di B , cioè $R \subseteq A \times B$, consideriamo una *nuova* relazione, che indicheremo con R^{-1} , tra gli elementi degli insiemi B e A , cioè $R^{-1} \subseteq B \times A$, ottenuta invertendo l'ordine di tutte le coppie di R . Tale nuova relazione R^{-1} è denominata *relazione inversa* della relazione R .

Definizione 2.2. *Data la relazione $R \subseteq A \times B$, si dice relazione inversa della R la relazione $R^{-1} \subseteq B \times A$ tale che $R^{-1} = \{(b, a) : (a, b) \in R\}$.*

Esempio 2.2. Si considerino gli insiemi $A = \{1, 2, 5\}$ e $B = \{1, 3\}$, e la relazione $R \subseteq A \times B$ tale che $R = \{(a, b) : a \leq b\} = \{(1, 1), (1, 3), (2, 3)\}$. La relazione inversa R^{-1} è $\{(b, a) : b \geq a\} = \{(1, 1), (3, 1), (3, 2)\}$.

2.2 Relazioni tra un insieme e se stesso e loro proprietà

Per definizione del prodotto cartesiano di due insiemi, non è necessario che gli insiemi in questione siano distinti: come abbiamo visto, è possibile parlare del prodotto cartesiano $I \times I$, costituito dall'insieme di tutte le coppie ordinate (a, b) con $a \in I$ e $b \in I$. Esamineremo ora le caratteristiche delle relazioni definite in $I \times I$; tali relazioni possono infatti godere di interessanti proprietà, introdotte dalle definizioni seguenti.

Definizione 2.3. Si dice che la relazione $R \subseteq I \times I$ gode della proprietà riflessiva se, per ogni $a \in I$, vale $(a, a) \in R$.

Cioè, una relazione definita tra gli elementi di un insieme gode della proprietà riflessiva (si dice anche “è riflessiva”) se *ogni* elemento dell’insieme considerato è in relazione con se stesso. Nella rappresentazione tabellare della relazione, questo significa che sono presenti gli elementi della diagonale.

Definizione 2.4. Si dice che la relazione $R \subseteq I \times I$ gode della proprietà

antiriflessiva se, per ogni $a \in I$, vale $(a, a) \notin R$.

Cioè, una relazione definita tra gli elementi di un insieme gode della proprietà antiriflessiva (si dice anche “è antiriflessiva”) se *nessun* elemento

dell'insieme considerato è in relazione con se stesso. Nella rappresentazione tabellare, questo significa che non è presente alcun elemento della diagonale.

Definizione 2.5. *Si dice che la relazione $R \subseteq I \times I$ gode della proprietà simmetrica se, per ogni $a, b \in I$, si ha che, se $(a, b) \in R$, allora $(b, a) \in R$.*

Cioè, una relazione definita tra gli elementi di un insieme gode della proprietà simmetrica (si dice anche “è simmetrica”) quando, per *ogni* coppia di elementi a e b dell'insieme considerato, accade che, se a è in relazione con b , allora anche b è in relazione con a . Nella rappresentazione tabellare, questo significa che la matrice è simmetrica.

Definizione 2.6. *Si dice che la relazione $R \subseteq I \times I$ gode della proprietà*

antisimmetrica se, per ogni $a, b \in I$, si ha che, se $(a, b) \in R$ e $(b, a) \in R$, allora $a = b$.

Cioè, una relazione definita tra gli elementi di un insieme gode della proprietà antisimmetrica (si dice anche “è antisimmetrica”) quando, per ogni coppia di elementi a e b dell’insieme considerato, il contemporaneo essere a in relazione con b e b in relazione con a implica che $a = b$.

Definizione 2.7. Si dice che la relazione $R \subseteq I \times I$ gode della proprietà transitiva se, per ogni $a, b, c \in I$, si ha che, se $(a, b) \in R$ e $(b, c) \in R$, allora $(a, c) \in R$.

Cioè, una relazione definita tra gli elementi di un insieme gode della proprietà transitiva (si dice anche “è transitiva”) quando, per ogni terna di

elementi a , b e c dell'insieme considerato, accade che, se a è in relazione con b e b è in relazione con c , allora a è in relazione con c .

Esempio 2.3. Consideriamo nell'insieme I delle rette del piano la $R \subseteq I \times I$ tale che $R = \{(r, s) : r \text{ è coincidente o parallela a } s\}$. Tale relazione gode delle proprietà:

1. riflessiva, perché ogni retta è coincidente o parallela a se stessa (nel caso specifico, coincidente);
2. simmetrica, perché se la retta r è coincidente o parallela alla retta s , allora anche s è coincidente o parallela a r ;
3. transitiva, perché se la retta r è coincidente o parallela alla retta s e la retta s è coincidente o parallela alla retta t , allora la retta r risulta coincidente o parallela a t .

Il lettore verificherà che R non gode delle altre proprietà sopra esaminate (cioè antiriflessiva e antisimmetrica).

Esempio 2.4. Consideriamo, nell'insieme J dei segmenti del piano, la $S \subseteq J \times J$ tale che $S = \{(a, b) : \text{la lunghezza di } a \text{ è non minore di quella di } b\}$. Tale relazione gode delle proprietà:

1. riflessiva, perché la lunghezza di ogni segmento è non minore di se stessa;
2. antisimmetrica, perché se la lunghezza di un primo segmento è non minore della lunghezza di un secondo ed inoltre la lunghezza del secondo segmento è non minore di quella del primo, allora i due segmenti considerati hanno la stessa lunghezza;
3. transitiva, perché se la lunghezza di un segmento a è non minore di quella di un segmento b e la lunghezza del segmento

b è non minore di quella di un segmento c, allora la lunghezza di a è non minore di quella di c.

La relazione S non gode delle altre proprietà sopra esaminate (antiriflessiva e simmetrica).

Esempio 2.5. Consideriamo, nell'insieme I delle rette del piano, la $T \subseteq I \times I$ tale che $T = \{(r, s) : r \text{ è perpendicolare a } s\}$. Tale relazione gode delle proprietà:

1. antiriflessiva, perché nessuna retta è perpendicolare a se stessa;
2. simmetrica, perché se una retta r è perpendicolare ad una retta s , allora la retta s è perpendicolare a r .

La relazione T non gode delle altre proprietà sopra esaminate (riflessiva,

antisimmetrica e transitiva).

Data una relazione non transitiva R , si può sempre costruire una “minima” relazione transitiva che la contenga come insieme; la denoteremo con il termine *chiusura transitiva* di R , scritta R^t . La chiusura transitiva di una relazione R può dunque essere intesa come l’intersezione di tutte le relazioni transitive che contengono R : così facendo, infatti, si otterrà una relazione (ancora transitiva) R^t contenente R , transitiva e con la proprietà di essere la più piccola possibile, appunto perché intersezione di tutte quelle soddisfacenti i requisiti. Ne consegue che, per ogni relazione R , la chiusura transitiva esiste ed è unica.

Alternativamente, la chiusura transitiva R^t di una relazione R può essere costruita considerando tutte le coppie (x, x') che

sono estremi di catene $(x, x_1, x_2, \dots, x_n, x)$ di cui ogni coppia di elementi consecutivi $(x, x_1), (x_1, x_2), \dots, (x_n, x)$ appartiene ad R . Che tale relazione sia transitiva è chiaro: se (x, x) e (x, x) appartengono ad Rt , allora esisterà una successione $(x, x_1), (x_1, x_2), \dots, (x_n, x'), \dots, (x_m, x'')$ che ci consentirà di affermare che anche $(x, x'') \in R^t$. Inoltre ogni relazione transitiva che contenga R dovrà contenere tutte le coppie estremi di catene finite come quelle descritte.

Esempio 2.6. Consideriamo l'insieme X dei punti di una città (il termine “punto” si intenda, in questo caso, informalmente, dunque più in senso “geografico” che geometrico). Sia A la relazione tra elementi di X individuata dalla proprietà “essere collegati da un autobus”. Lasciamo al lettore il compito di verificare che essa è non riflessiva e

simmetrica. Inoltre la relazione A non è transitiva in quanto può accadere che sia impossibile andare dal punto P al punto Q utilizzando un solo autobus pur potendo andare in questo modo da P a S e da S a Q.

Per costruire la chiusura transitiva A^t , cioè la minima relazione transitiva contenente A, dobbiamo “ampliare” A considerando in relazione anche punti di X raggiungibili tra di loro mediante un certo numero (finito, anche se non fissato) di autobus da prendere uno di seguito all’altro: e questa è proprio l’operazione che comunemente viene fatta da chi utilizza i mezzi pubblici nel programmare uno spostamento.

2.3 Relazioni di equivalenza e insieme quoziante

Fin dall'antichità, accanto all'operazione mentale di “raccolta di elementi” per formare un insieme (o classe), è stata considerata quella di “astrazione” rispetto a certe proprietà per formare individui di livello superiore. Ad esempio, se considero l’insieme degli animali, ma non mi interessano i singoli individui quanto, piuttosto, individui a meno di certe caratteristiche; si può procedere col definire una caratteristica importante (ad esempio la possibilità di essere fecondi con prole feconda) e considerare indifferenti due individui che siano indistinguibili rispetto a questa proprietà. A questo punto, non si ha più a che fare con gli animali, ma con le specie di animali, cioè con entità più astratte.

Questo procedimento molto generale per costruire nuovi insiemi, può essere formalizzato introducendo il concetto di

relazione di equivalenza, che permette di formulare la non-distinguibilità degli individui di un dato insieme.

Definizione 2.8. *Una relazione $R \subseteq I \times I$ si dice relazione di equivalenza se gode delle proprietà riflessiva, simmetrica e transitiva.*

Esempio 2.7. Consideriamo, nell'insieme I delle rette del piano, la relazione $R = \{(r, s) : r \text{ è coincidente o parallela a } s\}$, già esaminata nell'[Esempio 2.3](#). La relazione data è una relazione di equivalenza in quanto gode delle proprietà riflessiva, simmetrica e transitiva.

Definizione 2.9. *Si dice classe di equivalenza dell'elemento $a \in I$ rispetto alla relazione di equivalenza $R \subseteq I \times I$ il sottoinsieme di I costituito dagli elementi*

x tali che $(a, x) \in R$; tale insieme verrà denotato con $[a]_R$ o più semplicemente $[a]$ se R è chiara dal contesto.

Si noti che le classi di equivalenza sono non vuote (per la proprietà riflessiva delle relazioni di equivalenza, alla classe $[a]$ appartiene sempre l'elemento a) e a due a due disgiunte. Proviamo quest'ultima affermazione.

Proposizione 2.1. *Siano a e b due elementi di I , siano $[a]$ e $[b]$ le classi di equivalenza di a e di b rispetto alla relazione di equivalenza $R \subseteq I \times I$, e sia $[a] \neq [b]$; allora $[a]$ e $[b]$ sono disgiunte.*

Dimostrazione. Mostreremo che se esiste un elemento $x \in I$ tale che $x \in [a]$ e $x \in [b]$, allora deve essere $[a] = [b]$. Infatti, per definizione, $x \in [a]$ significa che $(a, x) \in R$ e $x \in [b]$ significa che $(b, x) \in R$. Per la

proprietà simmetrica e transitiva della relazione R , $(a, x) \in R$ e $(b, x) \in R$ implicano $(a, b) \in R$ e quindi $[a] = [b]$.

□

Definizione 2.10. Si dice insieme quoziante dell'insieme I rispetto alla relazione R l'insieme I / R delle classi di equivalenza degli elementi di I rispetto alla relazione di equivalenza R .

L'insieme quoziante costituisce una *partizione* di I : con ciò intendiamo che l'insieme quoziante è un insieme di sottoinsiemi di I non vuoti, a due a due disgiunti (come sopra dimostrato) e la cui l'unione è I .

Esempio 2.8. Consideriamo, nell'insieme I delle rette del piano, la relazione di equivalenza $R = \{(r, s) : r \text{ è coincidente o parallela a } s\}$ già esaminata

precedentemente. Le classi di equivalenza rispetto a tale relazione sono i sottoinsiemi di I del tipo $[s] = \{r \in I : r \text{ è coincidente o parallela a } s\}$. Ciascuno degli elementi dell'insieme quoziante I / R è costituito da un insieme di rette parallele: può essere identificato con la comune direzione di esse.

Di fatto l'uguaglianza di elementi in un insieme matematico è praticamente sempre definita attraverso una relazione di equivalenza ed un quoziante da un insieme meno astratto. Si pensi per esempio a come i numeri interi \mathbb{Z} vengano ottenuti dai naturali \mathbb{N} (si veda l'[Esercizio 2.24](#)) o come i numeri razionali \mathbb{Q} vengano ottenuti dagli interi \mathbb{Z} (si veda l'[Esercizio 2.13](#)).

2.4 Relazioni d'ordine

Introduciamo ora un'altra importante classe di relazioni.

Definizione 2.11. Una relazione $R \subseteq I \times I$ si dice relazione d'ordine se gode delle proprietà riflessiva, antisimmetrica e transitiva.

Definizione 2.12. Una relazione $R \subseteq I \times I$ si dice relazione di ordine stretto se gode delle proprietà antiriflessiva e transitiva.

Esempio 2.9. Consideriamo nell'insieme J dei segmenti del piano la $S \subseteq J \times J$ tale che $S = \{(a, b) : \text{la lunghezza di } a \text{ è non minore di quella di } b\}$, già esaminata nell'[Esempio 2.4](#). La relazione data è una relazione d'ordine (largo) in quanto gode delle proprietà riflessiva, antisimmetrica e transitiva.

Esempio 2.10. Consideriamo nell'insieme J dei segmenti del piano la $U \subseteq J \times J$ tale che $U = \{(a, b) : \text{la lunghezza di } a \text{ è maggiore di quella di } b\}$. La relazione data gode delle proprietà:

1. antiriflessiva, perché nessun segmento può avere lunghezza maggiore della propria stessa lunghezza;
2. transitiva, perché se un segmento a ha lunghezza maggiore di quella di un segmento b ed il segmento b ha lunghezza maggiore di quella di un segmento c , allora a ha lunghezza maggiore di c .

La relazione U è una relazione d'ordine stretto; inoltre, non gode delle altre proprietà esaminate nel paragrafo precedente.

Siamo abituati a considerare, per ogni

relazione d'ordine largo, la corrispondente relazione di ordine stretto: si pensi a \leq e $<$ tra numeri. Questa situazione è del tutto generale: data una relazione d'ordine largo, possiamo sempre ottenere una di ordine stretto togliendo tutte le coppie del tipo (x, x) ; viceversa, data una relazione d'ordine stretto, possiamo sempre ottenere una di ordine largo aggiungendo tutte le coppie del tipo (x, x) . Si noti anche che l'inversa di una relazione d'ordine (ad esempio, \geq è l'inversa di \leq) è ancora una relazione d'ordine.

Le relazioni d'ordine (largo o stretto) definite tra gli elementi di un insieme I servono a *confrontare* due elementi di I , stabilendo quale dei due elementi in questione preceda l'altro in una “classifica” basata sull'ordinamento indotto dalla relazione. Ad esempio, occupiamoci della relazione d'ordine largo

$S = \{(a, b) : \text{la lunghezza di } a \text{ è non minore di quella di } b\}$ introdotta sull'insieme J dei segmenti del piano, precedentemente esaminata: in base ad essa, è possibile “ordinare” l'insieme J rispetto alla lunghezza:

1. se $(a, b) \in S$, allora a precede b nell'ordinamento;
2. se $(b, a) \in S$, allora b precede a nell'ordinamento.

In tale caso, qualsiasi siano gli elementi a e b scelti in J , si verifica sempre *uno* dei due casi $(a, b) \in S$ oppure $(b, a) \in S$ (per segmenti non aventi la stessa lunghezza tali possibilità si escludono a vicenda): infatti, assegnati due segmenti a e b , accade sempre che

1. la lunghezza di a è non minore di quella di b , oppure

2. la lunghezza di b è non minore di quella di a

Ma accade *sempre* così? In altri termini: assegnata in un qualsiasi insieme una qualsiasi relazione d'ordine S , accade sempre che, detti a e b due qualsiasi elementi di tale insieme, si verifichi una delle due possibilità $(a, b) \in S$ o $(b, a) \in S$? Oppure può accadere che, per almeno una coppia di elementi c e d dell'insieme esaminato, risulti $(c, d) \notin S$ ed anche $(d, c) \notin S$? Queste ultime situazioni sono possibili: la distinzione ora presentata è collegata alla possibilità di confrontare tutte le coppie di elementi dell'insieme I in base alla relazione d'ordine assegnata in I e da ciò dipende la possibilità di ordinare tutti gli elementi di I in base a quanto considerato nella relazione.

Una relazione d'ordine S definita in un insieme I e tale che, per ogni coppia di

elementi a e b di I si verifica sempre uno dei casi $(a, b) \in S$ oppure $(b, a) \in S$ (ovvero che consenta il confronto di tutte le coppie di elementi di I) si dice *relazione d'ordine totale*; una relazione d'ordine S che non per tutte le coppie consenta tale confronto, ovvero tale che per almeno una coppia di elementi c e d dell'insieme considerato accada che $(c, d) \notin S$ e che $(d, c) \notin S$, si dice *relazione d'ordine parziale*.

Osservazione. Il problema, ora presentato nel caso di una relazione d'ordine largo, si può estendere alle relazioni d'ordine stretto, come la U presentata nell'[Esempio 2.10](#). Le possibilità, in quest'ultimo caso, non sono più due ma tre: a $(a, b) \in U$ e $(b, a) \in U$, corrispondenti ai casi in cui la lunghezza di a è maggiore di quella di b e viceversa, deve essere aggiunto il caso in cui le lunghezze sono uguali.

Esempio 2.11. Sia I l'insieme avente per elementi tutti i sottoinsiemi di \mathbb{R} (cioè l'insieme delle parti di \mathbb{R}); definiamo la relazione $S \subseteq I \times I$ tale che $S = \{(A, B) : A \subseteq B\}$. Tale relazione è una relazione d'ordine, in quanto gode delle proprietà:

1. riflessiva, perché per ogni A , vale che $A \subseteq A$;
2. antisimmetrica, perché da $A \subseteq B$ e $B \subseteq A$ segue che $A = B$;
3. transitiva, perché da $A \subseteq B$ e $B \subseteq C$ segue che $A \subseteq C$.

La S è una relazione d'ordine parziale: infatti è possibile trovare coppie di sottoinsiemi C e D di \mathbb{R} tali che $(C, D) \notin S$ e $(D, C) \notin S$. Ad esempio, se $C = \{x \in \mathbb{R} : 0 < x < 2\}$ e $D = \{x \in \mathbb{R} : 1 < x < 3\}$, non essendo $C \subseteq D$ né $D \subseteq C$, risulta $(C, D) \notin S$ e $(D, C) \notin S$.

Consideriamo ora un caso ancora meno rigido.

Esempio 2.12. Supponiamo di voler ordinare il popolo italiano in base al censimento. Sappiamo benissimo introdurre una relazione $R \subseteq I \times I$ tra gli italiani dicendo che $(a, b) \in R$ se a non è più ricco di b . R è riflessiva e transitiva, ma non antisimmetrica, perché due italiani possono essere ugualmente ricchi senza essere la stessa persona. Possiamo allora introdurre una relazione di equivalenza $U \subseteq I \times I$ contenente tutte e sole le coppie (a, b) tali che a non è più ricco di b e b non è più ricco di a , cioè $(a, b) \in R$ e $(b, a) \in R$. A questo punto sull'insieme quoziente di I rispetto ad U delle classi di reddito degli italiani, possiamo facilmente definire una R' in modo analogo ad R ed ottenere una relazione d'ordine.

Definizione 2.13. Una relazione $R \subseteq I \times I$ si dice relazione di preordine se gode delle proprietà riflessiva e transitiva.

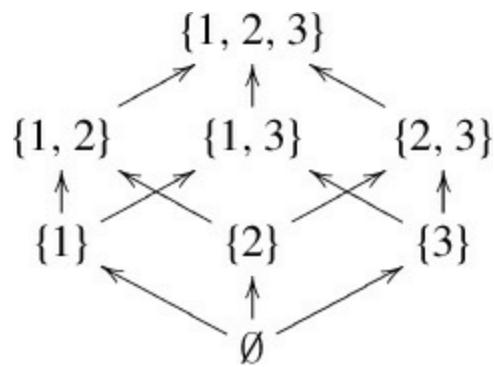
Proposizione 2.2. Data una relazione $R \subseteq I \times I$ di preordine si può definire un ordinamento corrispondente R' sull'insieme I' , quoziante di I rispetto alla relazione di equivalenza $U = R \cap R^{-1}$.

Infine, notiamo che a volte un insieme ordinato può essere specificato in maniera diagrammatica, collegando con una freccia il primo elemento di ogni coppia della relazione con il rispettivo secondo elemento. Tuttavia, per evitare di complicare troppo il diagramma della relazione R , spesso si disegna il diagramma della sotto-relazione $S \subseteq R$ tale che la sua chiusura transitiva S^t coincida con R . Ad esempio, l'insieme dei naturali, ordinato secondo la relazione \leq ,

può essere scritto come

$$0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow \dots$$

mentre l'insieme dei sottoinsiemi di $\{1, 2, 3\}$, ordinato mediante inclusione insiemistica, può essere scritto come



Infine, rispetto ad una relazione d'ordine, possono esistere in un insieme degli elementi “speciali”.

Definizione 2.14.. *Dato un insieme X e una relazione d'ordine \leq su di esso, si dice:*

- massimo *un elemento* x_1 t.c. per ogni $x \in X$ accade $x \leq x_1$;
- minimo *un elemento* x_0 t.c. per ogni $x \in X$ accade $x_0 \leq x$;
- massimo comune minorante *tra due elementi* $x, y \in X$ un elemento $z \in X$ t.c. $z \leq x$ e $z \leq y$, e per ogni $z' \in X$ t.c. $z' \leq x$ e $z' \leq y$ accade $z' \leq z$;
- minimo comune maggiorante *tra due elementi* $x, y \in X$ un elemento $t \in X$ t.c. $x \leq t$ e $y \leq t$, e per ogni $t' \in X$ t.c. $x \leq t'$ e $y \leq t'$ accade $t \leq t'$.

Chiaramente, tali elementi non esistono sempre, ma si può provare come facile esercizio che, se esistono, allora sono unici.

Esempio 2.13. Si consideri l'insieme dei numeri naturali positivi con la relazione di divisibilità: esiste un minimo (il numero 1), non esiste un massimo,

esistono massimo comune minorante e minimo comune maggiorante di due numeri e sono rispettivamente il loro massimo comun divisore (M.C.D.) e minimo comune multiplo (m.c.m.).

Esempio 2.14. *Si consideri l'insieme dei sottoinsiemi di un insieme A con la relazione di inclusione: esiste un minimo (l'insieme \emptyset), esiste un massimo (A), esistono massimo comune minorante e minimo comune maggiorante di sottoinsiemi e sono rispettivamente la loro intersezione e la loro unione.*

Come già notato, la relazione inversa di una relazione d'ordine è ancora d'ordine. Se dunque invertiamo la relazione, il minimo diventerà il massimo e viceversa, mentre il massimo comune minorante diventerà il minimo comune maggiorante e viceversa. Così tutto ciò che potremo

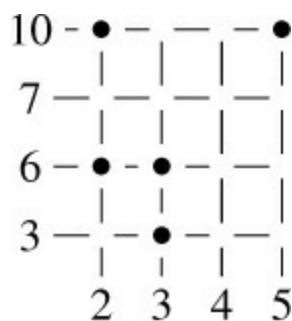
dire per una di queste nozioni, potremo dirla, *mutatis mutandis*, anche per l'altra. Incontriamo qui per la prima volta la *dualità*, cioè quella proprietà di “simmetria” di alcune strutture matematiche, per la quale esiste una trasformazione della struttura in sé stessa che permette di ritrovare una parte delle nozioni ad essa relative in un’altra parte dopo la trasformazione. La nuova struttura non è necessariamente coincidente con la prima, ma è dello stesso tipo. In questi casi il *principio di dualità* permetterà di dimezzare definizioni e dimostrazioni relative alla generica struttura di quel tipo. Un esempio ancora più stringente di dualità sarà realizzato dalle algebre di Boole che vedremo più avanti.

Esercizi svolti

Esercizio 2.1. Sia R la relazione tra $E = \{2, 3, 4, 5\}$ e $F = \{3, 6, 7, 10\}$ definita dalla proposizione “ x divide y ”. Scrivere R come insieme di coppie ordinate e disegnare R sul diagramma delle coordinate di $E \times F$.

Soluzione: Ricordiamo che “ x divide y ” vuol dire che y è un multiplo (non necessariamente proprio) di x , cioè $y = x \cdot k$ per qualche $k \in \mathbb{N}_0$ (si noti quindi che qualunque numero divide se stesso).

Pertanto, $R = \{(2, 6), (2, 10), (3, 3), (3, 6), (5, 10)\}$ che, scritta in forma diagrammatica, è:



Esercizio 2.2. Ciascuna delle seguenti proposizioni definisce una relazione su \mathbb{N} . Dire per ciascuna relazione se essa è riflessiva:

- (a) “ x è minore o uguale a y ”
- (b) “ x divide y ”
- (c) “ $x + y = 10$ ”
- (d) “ x e y sono primi tra loro”

Soluzione: Chiaramente, ogni numero è minore o uguale a sé (in particolare, è uguale); quindi la relazione del punto (a) è riflessiva. Abbiamo già detto che ogni numero divide se stesso: anche la (b) è riflessiva. La (c) non definisce una relazione riflessiva: per esempio, prendendo $x = 0$, abbiamo che $x + x = 0 = 10$. Infine, ricordiamo che due numeri sono primi tra loro se il loro unico divisore comune è 1; pertanto, la (d) non è una

relazione riflessiva, poiché il generico numero x ha in comune con sé tutti i suoi divisori (oltre 1 c'è sempre anche x stesso).

Esercizio 2.3. *Si considerino le relazioni definite nell'Esercizio 2.2 e si dica se sono antiriflessive o meno.*

Soluzione: Nessuna è antiriflessiva.

Infatti, la relazione del punto (a) non è antiriflessiva poiché ogni numero è minore o uguale a sé (N.B.: le cose sarebbero diverse se invece di ' \leq ' avessimo considerato ' $<$ '); la relazione del punto (b) non è antiriflessiva poiché ogni numero divide se stesso; la relazione del punto (c) non è antiriflessiva poiché $5 + 5 = 10$; la relazione del punto (d) non è antiriflessiva poiché l'unico divisore comune ad 1 e se stesso è 1.

Esercizio 2.4. *Una relazione definita su*

un insieme A può essere sia riflessiva che antiriflessiva?

Soluzione: Sì, ma soltanto se A è l'insieme vuoto.

Esercizio 2.5. *Si considerino le relazioni definite nell'[Esercizio 2.2](#) e si dica se sono simmetriche o meno.*

Soluzione: La relazione del punto (a) non è simmetrica: $3 \leq 5$ ma $5 > 3$. Anche la relazione del punto (b) non è simmetrica: 2 divide 4, mentre 4 non divide 2. La (c) definisce una relazione simmetrica: se $x + y = 10$ allora, per la proprietà commutativa della somma, $y + x = 10$. Infine, anche la (d) è una relazione simmetrica: i divisori comuni di x e y sono i divisori comuni di y e x .

Esercizio 2.6. *Esiste un insieme A tale*

che qualsiasi relazione su A è simmetrica?

Soluzione: Sì, per esempio $A = \{a\}$ (per un generico elemento a) oppure $A = \emptyset$. Nel secondo caso, l'unica relazione su A è la relazione vuota; visto che tale relazione non contiene alcuna coppia, la proprietà di simmetria è banalmente verificata. Nel primo caso, invece, $A \times A = \{(a, a)\}$; pertanto, le uniche due relazioni su A sono la relazione vuota e $A \times A$, entrambe simmetriche.

Esercizio 2.7. *Si considerino le relazioni definite nell'[Esercizio 2.2](#) e si dica se sono antisimmetriche o meno.*

Soluzione: La relazione del punto (a) è antisimmetrica: l'unica possibilità affinché sia $x \leq y$ che $y \leq x$ siano vere è che $x = y$. Anche la relazione del punto (b) è antisimmetrica: se x divide y e viceversa,

deve essere $x = y$. La (c) non definisce una relazione antisimmetrica: $3 + 7 = 10$, $7 + 3 = 10$ ma $3 \neq 7$. Infine, anche la (d) non è una relazione antisimmetrica: 3 e 5 sono primi tra loro, 5 e 3 sono primi tra loro, ma sono diversi.

Esercizio 2.8. *Una relazione R definita su un insieme A può essere sia simmetrica che antisimmetrica?*

Soluzione: Sì, per esempio l'identità su A o la relazione vuota su A .

Esercizio 2.9. *Si considerino le relazioni definite nell'[Esercizio 2.2](#) e si dica se sono transitive o meno.*

Soluzione: La relazione del punto (a) è banalmente transitiva: se x è al più y che è al più z , allora x è al più z . Anche la relazione del punto (b) è transitiva: se x

divide y e y divide z (cioè $y = x \cdot k$ e $z = y \cdot h$), allora x divide z (poiché $z = y \cdot h = (x \cdot k) \cdot h = x \cdot (k \cdot h)$). La (c) non definisce una relazione transitiva: $3 + 7 = 10$ e $7 + 3 \neq 10$, mentre $3 + 3 = 10$. Infine, anche la (d) non è una relazione transitiva: 3 e 5 sono primi tra loro, 5 e 6 sono primi tra loro, ma 3 e 6 non sono primi tra loro (ammettono come divisori comuni 1 e 3).

Esercizio 2.10. Si consideri l'insieme $I = \{Milano, Napoli, Palermo, Varese, Caserta\}$. Definire una o più relazioni di equivalenza su I e, per ciascuna relazione R , determinare l'insieme quoziante I / R .

Soluzione: Una possibile relazione di equivalenza è “essere nella stessa regione di”; per una tale R , si ha che $I / R = \{\{Milano, Varese\}, \{Palermo\}, \{Napoli, Caserta\}\}$. Altre possibili relazioni d'equivalenza potrebbero essere: “avere la

stessa targa automobilistica di” oppure “essere nello stesso stato di”. Nel primo caso, l’insieme quoziente coinciderebbe con la partizione che contiene cinque blocchi di un solo elemento; nel secondo caso, avremmo invece la partizione formata da un unico blocco che contiene tutte e cinque le città.

Esercizio 2.11. *Si consideri l’insieme di tutte le possibili coppie di numeri interi, il cui secondo elemento non sia nullo, e si interpreti la coppia (n, m) come la frazione $\frac{n}{m}$. Si costruisca una partizione dell’insieme considerato tale che ciascuna classe di equivalenza contenga coppie equivalenti, del tipo $\frac{ad}{de} = \frac{bc}{cf}, \dots$. Qual’è la relazione di equivalenza che genera tale partizione? A quale tipo di numero corrisponde ciascuna partizione?*

Soluzione: La partizione richiesta può

essere ottenuta mediante il quoziente mediante la relazione d'equivalenze \sim tale che $(m, n) \sim (p, q)$ se e soltanto se esistono h e k non nulli tali che $mh = pk$ e $nh = qk$. Tutti gli elementi di uno stesso blocco hanno in comune il fatto di rappresentare frazioni che, una volta semplificate, si riducono alla stessa frazione. Inoltre, ogni blocco conterrà una coppia che rappresenta una frazione non semplificabile; tale coppia può essere considerata il rappresentante della classe.

Esercizio 2.12. Si consideri l'insieme $\mathbb{Z} \times \mathbb{Z}_0$, cioè l'insieme di coppie ordinate dei numeri interi, la cui seconda componente non sia nulla. Sia \simeq la relazione su $\mathbb{Z} \times \mathbb{Z}_0$ definita nel modo seguente: (a, b) è in relazione \simeq (c, d) (scritto $(a, b) \simeq (c, d)$) se e solo se $ad = bc$. Dimostrare che \simeq è una relazione di equivalenza.

Soluzione: Dobbiamo dimostrare che ‘ \simeq ’ gode delle proprietà riflessiva, simmetrica e transitiva.

- Per ogni $(a, b) \in \mathbb{Z} \times \mathbb{Z}_0$ si ha che $(a, b) \simeq (a, b)$ poiché $ab = ba$.
- Se $(a, b) \simeq (c, d)$ allora $(c, d) \simeq (a, b)$ poiché $ad = bc$ implica $cb = da$.
- Se $(a, b) \simeq (c, d)$ e $(c, d) \simeq (e, f)$ allora $(a, b) \simeq (e, f)$ poiché $ad = bc$ e $cf = de$ implicano $af = be$. Infatti:
 - se $c, e \neq 0$ allora $\frac{a}{e} = \frac{b}{f}$ da cui $\frac{a}{e} = \frac{b}{f}$ da cui $af = be$;
 - se $c = 0$ allora $a = 0$ ed $e = 0$ (infatti, per definizione, $d, f \neq 0$), e questo banalmente implica $af = be$;
 - se $e = 0$ la prova è analoga al caso precedente.

Esercizio 2.13. In che modo gli [Esercizi 2.11](#) e [2.12](#) sono collegati l'uno all'altro?

Soluzione: Definiscono la stessa relazione; andiamo infatti a dimostrare che le relazioni \sim (la relazione definita nell'[Esercizio 2.11](#)) e \simeq (la relazione definita nell'[Esercizio 2.12](#)) coincidono (come insiemi di coppie).

- $\sim \subseteq$. Sia $(a, b) \sim (c, d)$. Per definizione, esistono due naturali h e k tali che $ah = ck$ e $bh = dk$. Ma allora $ahd = bhck$, da cui $ad = bc$.
- $\subseteq \sim$. Sia $(a, b) \simeq (c, d)$; per definizione $a \cdot d = b \cdot c$. Poiché $b, d \neq 0$, possiamo prendere come h e k i numeri d e b , rispettivamente: infatti sono non nulli, $ad = cb$ e $bd = db$. Quindi $((a, b), (c, d)) \in \sim$.

Pertanto, $(\mathbb{Z} \times \mathbb{Z}_0)/\simeq = (\mathbb{Z} \times \mathbb{Z}_0)/\sim$; inoltre, è facile convincersi che tali insiemi quoziensi possono essere usati per definire i numeri razionali \mathbb{Q} (si veda il modo in cui

abbiamo definito il rappresentante di ogni classe nell'[Esercizio 2.11](#)).

Esercizio 2.14. Sia $B = \{a, b, c, d, e\}$ ordinato come segue:

$$\aleph_0$$

Sia B la famiglia di tutti i sottoinsiemi totalmente ordinati di B che contengono due o più elementi, e sia B parzialmente ordinato mediante inclusione insiemistica: costruire un diagramma dell'ordinamento di B .

Soluzione: L'insieme B contiene gli insiemi totalmente ordinati $\{d, b\}$, $\{d, a\}$, $\{d, b, a\}$, $\{e, b\}$, $\{e, a\}$, $\{e, b, a\}$, $\{e, c\}$, $\{e, c, a\}$, $\{b, a\}$ e $\{c, a\}$. L'ordinamento parziale di B tramite inclusione insiemistica è il seguente:

$$2^{\aleph_0}$$

Esercizi da svolgere

Esercizio 2.15. Sia R la relazione tra $E = \{2, 3, 4, 5\}$ e $F = \{3, 4, 5, 6\}$ definita dalla proposizione “ x è minore di y ”. Scrivere R sia come insieme di coppie che in maniera diagrammatica.

Esercizio 2.16. Sia $C = \{1, 2, 3\}$ e si considerino le seguenti relazioni su C :

- (a) $\{(1, 2), (3, 2), (2, 2), (2, 3)\}$
- (b) $\{(1, 2)\}$
- (c) $\{(1, 2), (2, 3), (1, 3)\}$
- (d) $\{(1, 1), (2, 2), (2, 3), (3, 2), (3, 3)\}$
- (e) $C \times C$

Per ognuna di esse, dire se è riflessiva, antiriflessiva, simmetrica, antisimmetrica e transitiva, giustificando la risposta.

Esercizio 2.17. Una relazione definita

su un insieme A può essere sia non riflessiva che non antiriflessiva?

Esercizio 2.18. *Si elenchino tutte le proprietà di cui godono le seguenti relazioni:*

1. *la relazione che accoppia circonferenze del piano concentriche;*
2. *la relazione che accoppia solidi con lo stesso volume;*
3. *la relazione di inclusione insiemistica;*
4. *la relazione “essere fratello di”;*
5. *la relazione “essere parallela a” (nel dominio delle rette del piano);*
6. *la relazione “avere la stessa area di” (tra figure del piano);*
7. *la relazione “essere perpendicolare a” (nel dominio delle rette del piano);*
8. *la relazione “essere divisore di” (tra numeri naturali).*

Esercizio 2.19. Sia $A = \{a, b, c, d, e, f, g\}$; dire se ciascuna delle seguenti famiglie di insiemi è una partizione di A :

1. $\{B_1 = \{a, c, e\}, B_2 = \{b\}, B_3 = \{d, g\}\}$;
2. $\{C_1 = \{a, e, g\}, C_2 = \{c, d\}, C_3 = \{b, f\}\}$;
3. $\{D_1 = \{a, b, e, g\}, D_2 = \{c\}, D_3 = \{d, f\}\}$;
4. $\{E_1 = \{a, b, c, d, e, f, g\}\}$.

Esercizio 2.20. Sia $S = \{a, b, c, d, e, f, g, h, i\}$ un insieme di studenti, di cui: a, d, e frequentano le lezioni del I canale; b, f, g, h frequentano il II canale; c, i frequentano il III canale. Si consideri la relazione $R(x, y) = "x$ frequenta lo stesso canale di $y"$. Scrivere R in rappresentazione caratteristica; dire se è possibile ottenere una partizione di S in classi di equivalenza secondo R e, in caso affermativo, scrivere tale partizione.

Esercizio 2.21. Sia $R(x, y)$ la relazione su \mathbb{Z} definita da “ $(x - y)$ è divisibile per 5”. Scrivere R in rappresentazione caratteristica e dimostrare che è una relazione di equivalenza.

Esercizio 2.22. Delle relazioni definite nell’Esercizio 2.18, si dica quali sono equivalenze, quali di ordine stretto e quali di ordine largo.

Esercizio 2.23. Sia A la famiglia di tutti i sottoinsiemi A dei numeri naturali, dove A ha le seguenti proprietà: A è finito e il massimo comune divisore degli elementi di A è 1.

1. Dire se ciascuno dei seguenti sottoinsiemi di \mathbb{N} appartiene ad A
 - (a) $\{3, 4, 8\}$
 - (b) $\{2, 3, 4, 8\}$
 - (c) $\{2, 4\}$

(d) $\{4, 5, 6, \dots\}$

(e) $\{4, 6, 8\}$

(f) $\{2, 3\}$

2. sia B la sottofamiglia di A che contiene gli insiemi, fra quelli elencati al punto precedente, che appartengono ad A , e si consideri un ordinamento su B definito dall'inclusione insiemistica: costruire un diagramma dell'ordinamento di B .

Esercizio 2.24. Similmente agli Esercizi 2.12 e 2.13, si definisca una relazione $\simeq \subseteq (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N})$ tale che $(\mathbb{N} \times \mathbb{N})/\simeq$ possa essere usata per definire i numeri interi \mathbb{Z} .

3

Funzioni

3.1 La definizione di funzione

Tra le relazioni binarie, cioè su una coppia di insiemi, alcune hanno una particolare importanza. Le *funzioni* (o *applicazioni*) sono relazioni tra gli elementi di un insieme D e di un insieme C tali che *ad ogni* elemento dell'insieme D corrisponda esattamente *un* (ovvero, *uno ed un solo*) elemento di C . In altre parole, una corrispondenza tra gli elementi di D e quelli di C è una funzione quando

1. ogni elemento di D ha un corrispondente in C , e
2. nessun elemento di D ha più di un corrispondente in C .

Definizione 3.1. *Una relazione $R \subseteq D \times C$ si dice funzione (o applicazione) se per ogni $a \in D$ esiste uno ed un solo $b \in C$ tale che $(a, b) \in R$.*

Una funzione si indica spesso con la lettera f e si scrive $f \subseteq D \times C$ o, più spesso, $f: D \rightarrow C$; il primo insieme, D , si dice *dominio* di f , il secondo, C , *codominio*. A ciascun elemento x del dominio della f corrisponde un'(unica) *immagine*, indicata da $f(x) \in C$. Analogamente, si dice che $x \in D$ è *controimmagine* di $f(x) \in C$.

Nota 3.1. *Se $f \subseteq D \times C$ è una funzione e se $b_1 \in C$ e $b_2 \in C$, con $b_1 = b_2$, in base*

alla definizione data risulta che

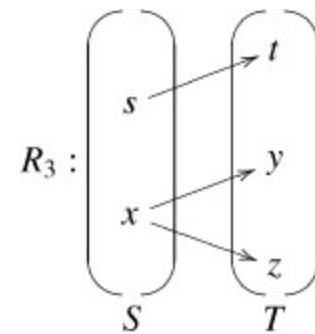
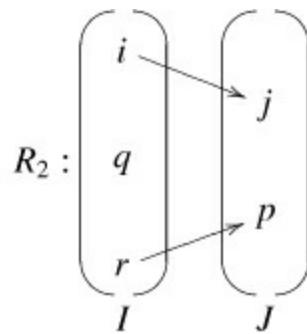
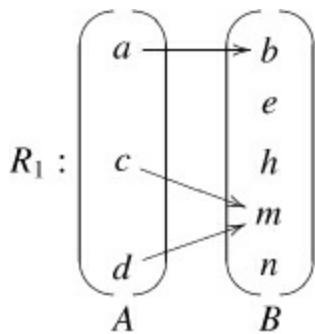
- se $(x, b_1) \in f$ allora $(x, b_2) \notin f$;
- se $(x, b_2) \in f$ allora $(x, b_1) \notin f$.

È invece possibile che esistano $b \in C$, $x_1 \in D$ e $x_2 \in D$, con $x_1 \neq x_2$ tali che $(x_1, b) \in f$ e $(x_2, b) \in f$.

Definizione 3.2. *Data la funzione $f: D \rightarrow C$, l'insieme $I m(D) \subseteq C$ costituito dalle immagini degli elementi di D si dice insieme delle immagini (o immagine) di f .*

Pertanto l'insieme delle immagini della funzione $f: D \rightarrow C$ è quel sottoinsieme di C costituito dagli elementi di C aventi (almeno) una controimmagine in D .

Esempio 3.1. *Consideriamo le relazioni rappresentate da:*



La prima di esse è una funzione, mentre le altre due non rispettano la definizione di funzione. Infatti, in $R_1 \subseteq A \times B$, ad ogni elemento di A corrisponde una ed una sola immagine in B ; l'insieme delle immagini della funzione R_1 è $\{b, m\}$. Nella relazione $R_2 \subseteq I \times J$, a $q \in I$ non corrisponde alcun elemento in J , contro la definizione di funzione. Nella relazione $R_3 \subseteq S \times T$, allo stesso elemento $x \in S$ corrispondono i distinti elementi $y \in T$ e $z \in T$, contro la definizione di funzione.

Esempio 3.2. Sia P l'insieme dei punti del piano e sia Z l'insieme delle circonferenze tracciate nel piano.

Consideriamo le relazioni $R \subseteq P \times Z$ e $S \subseteq Z \times P$:

- $R = \{(p, z) : il\ punto\ p \in P\ è\ il\ centro\ della\ circonferenza\ z \in Z\}$;
- $S = \{(z, p) : la\ circonferenza\ z \in Z\ ha\ per\ centro\ il\ punto\ p \in P\}$.

Esaminiamo la R : in essa, ad ogni punto del piano corrispondono infinite circonferenze, in quanto ogni punto può essere centro di infinite circonferenze (concentriche): pertanto, la R non rispetta la definizione di funzione. Nel caso della S , invece, ad ogni circonferenza z corrisponde (in qualità di centro) uno ed un solo punto: pertanto, la S è una funzione.

3.2 Funzioni iniettive, suiettive, biiettive

Nel paragrafo precedente abbiamo dato la definizione di funzione: abbiamo cioè precisato che una relazione tra due insiemi è detta funzione quando ad ogni elemento del primo insieme (dominio) corrisponde una ed un sola immagine nel secondo insieme.

Riflettiamo ora sulla funzione R_1 definita nell'[Esempio 3.1](#); dal suo esame, appare evidente che la definizione di funzione

1. *non* impone che due distinti elementi del dominio abbiano immagini distinte (a questo proposito si veda la Nota 3.1);
2. *non* impone che tutti gli elementi del secondo insieme abbiano una controimmagine nel dominio (ovvero che l'insieme delle immagini della funzione coincida con l'intero secondo insieme).

Quelle (particolari) funzioni che rispettano una di queste due ulteriori condizioni (oppure entrambe) vengono indicate con denominazioni specifiche.

Definizione 3.3. *La funzione $f: D \rightarrow C$ si dice iniettiva se per ogni $x_1 \in D$ e $x_2 \in D$, con $x_1 \neq x_2$, si ha che $f(x_1) \neq f(x_2)$.*

Dunque una funzione $f: D \rightarrow C$ si dice iniettiva se ad *ogni* coppia di elementi *distinti* di D corrisponde una coppia di elementi *distinti* di C , ovvero se nessun elemento di C è dotato di più di una controimmagine in D .

Definizione 3.4. *La funzione $f: D \rightarrow C$ si dice suriettiva se per ogni $b \in C$ esiste (almeno) un $x \in D$ tale che $f(x) = b$.*

Dunque una funzione $f: D \rightarrow C$ si dice suriettiva se *ogni* elemento dell'insieme C

ha (almeno) una controimmagine nel dominio, ovvero se l'insieme delle immagini di f coincide con *tutto* l'insieme C .

Definizione 3.5. *La funzione $f: D \rightarrow C$ si dice biiettiva (o biiezione o corrispondenza biunivoca) se è contemporaneamente iniettiva e suriettiva.*

Esempio 3.3. *La funzione $R_1: A \rightarrow B$ introdotta nell'[Esempio 3.1](#) non è iniettiva, in quanto ai due (distinti) elementi del dominio $c \in A$ e $d \in A$, corrisponde la stessa immagine $m \in B$. Essa non è suriettiva, in quanto gli elementi del secondo insieme e , h ed n non hanno alcuna controimmagine nel dominio A (cioè, l'insieme delle immagini della funzione, $\{b, m\}$, non coincide con tutto B). Non essendo iniettiva (né suriettiva),*

la funzione non è certamente biiettiva.

Esempio 3.4. Consideriamo la funzione S introdotta nell'[Esempio 3.2](#). Essa rispetta la definizione di funzione. Essa non è iniettiva: esistono coppie di circonferenze distinte aventi lo stesso centro (concentriche). È suriettiva: ogni punto del piano è centro di (almeno) una circonferenza (di infinite circonferenze, ma la precisazione è ininfluente per quanto riguarda la suriettività). Non è biiettiva, non essendo iniettiva.

Esempio 3.5. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ la funzione che ad ogni $x \in \mathbb{R}$ fa corrispondere il doppio di x , cioè $2x \in \mathbb{R}$. Si verifichi innanzitutto per esercizio che essa rispetta la definizione di funzione; inoltre, essa è una funzione iniettiva, suriettiva e biiettiva. Infatti f è iniettiva, in quanto ad ogni coppia di reali distinti

corrispondono coppie di doppi distinti: se le immagini 2a e 2b coincidessero, non potrebbero che coincidere anche a e b . La funzione f è inoltre suriettiva, in quanto ogni elemento del secondo insieme ha una controimmagine nel dominio: ovvero ogni reale y è il doppio di un (opportuno) reale, cioè $\frac{y}{2}$. Essendo iniettiva e suriettiva, f è biiettiva.

3.3 Composizione di funzioni

Consideriamo tre insiemi A , B e C , e le funzioni $f: A \rightarrow B$ e $g: B \rightarrow C$. La f fa corrispondere ad ogni $a \in A$ uno ed un solo $b \in B$; a tale elemento b , la g fa corrispondere uno ed un solo $c \in C$.

Possiamo riassumere quanto detto nello schema $a \xrightarrow{f} b \xrightarrow{g} c$, ovvero, con riferimento all'elemento $a \in A$ da cui trae origine la corrispondenza, $a \xrightarrow{f} f(a) \xrightarrow{g} g(f(a))$. È possibile

considerare una *nuova* funzione che faccia direttamente corrispondere all'elemento $a \in A$ l'elemento $g(f(a)) \in C$.

Definizione 3.6. *Date le funzioni $f: A \rightarrow B$ e $g: B \rightarrow C$, si dice funzione composta la funzione $g \circ f$ (*a volte scritta semplicemente gf*) che ad ogni elemento $a \in A$ fa corrispondere l'elemento $g(f(a)) \in C$.*

Attenzione: per convenzione, si indica *per prima la funzione che opera per ultima!* Ciò viene scelto in analogia con la scrittura $g(f(a))$ dell'elemento corrispondente di a nella funzione composta $g \circ f$.

Esempio 3.6. *Siano date le funzioni $f: \mathbb{R} \rightarrow \mathbb{R}$ e $g: \mathbb{R} \rightarrow \mathbb{R}$ definite da $f(x) = 3x$ e $g(x) = x + 2$; determiniamo le funzioni*

composte $f \circ g$ e $g \circ f$. Ricordiamo che, nelle funzioni composte, la funzione componente ad operare per prima è quella scritta per seconda. Iniziamo quindi con il ricavo dell'espressione di $f \circ g$:

$$f \circ g: x \rightarrow x + 2 \rightarrow 3(x + 2)$$

Osserviamo che la f , che all'elemento del dominio fa corrispondere il suo triplo, non opera sulla x , bensì sull'elemento $(x + 2)$, già trasformato dalla precedente azione della funzione g . Per quanto riguarda la $g \circ f$, otteniamo:

$$g \circ f: x \rightarrow 3x \rightarrow 3x + 2$$

Confrontando le funzioni $f \circ g$ e $g \circ f$ ricavate nell'esempio precedente, possiamo notare che la composizione di

funzioni *non gode della proprietà commutativa* anche se dominio e codominio sono lo stesso insieme. Si verifica invece che la composizione di funzioni *gode della proprietà associativa* non appena le funzioni considerate siano componibili; ovvero, assegnate le tre funzioni componibili f , g ed h , risulta $(f \circ g) \circ h = f \circ (g \circ h)$.

Proposizione 3.1. *Se $g \circ f$ è definita e sia g che f sono iniettive (rispettivamente suriettive o biunivoche), allora anche $g \circ f$ è iniettiva (rispettivamente suriettiva).*

3.4 Funzioni parziali

La definizione di funzione richiede che, per definire una funzione $f: D \rightarrow C$, siano specificati un insieme D , il dominio, un

secondo insieme C , il codominio, ed una legge che ad ogni $x \in D$ faccia corrispondere uno ed un solo $f(x) \in C$. Spesso, però, si considerano funzioni dette *parziali*, cioè ‘apparentemente’ definite sul loro dominio, ad esempio da \mathbb{R} ad \mathbb{R} , ma in realtà effettivamente definite soltanto da un sottoinsieme D di \mathbb{R} ed aventi immagini reali. Tale sottinsieme viene detto *dominio di definizione*.

Esempio 3.7. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ tale che $f(x) = \sqrt{x}$; allora il più grande $D \subseteq \mathbb{R}$ tale che per ogni $x \in D$ sia possibile calcolare $f(x) \in \mathbb{R}$ è $\{x \in \mathbb{R} : x \geq 0\}$. Dunque si dice che il dominio (di definizione) della f è $\{x \in \mathbb{R} : x \geq 0\}$.

In questi casi frequentemente la determinazione del *dominio di definizione* viene lasciata come esercizio. Questo modo di procedere è spesso accettato: è

però opportuno *specificare esplicitamente il dominio nei casi in cui possano sorgere malintesi*. Naturalmente possiamo sempre *restringere la definizione di una funzione parziale al suo dominio di definizione, rendendola così totale*.

Esempio 3.8. Consideriamo la funzione f espressa da:

$$f(x) = \frac{x}{\sqrt{x-1}}$$

Se cerchiamo un dominio $D \subseteq \mathbb{R}$ in modo che il denominatore sia non nullo e la radice quadrata sia reale, dobbiamo imporre la condizione $x > 1$. Il discorso sarebbe diverso se fossimo partiti non da \mathbb{R} ma da \mathbb{C} : infatti nei numeri complessi ogni valore $x \neq 0$ sarebbe accettabile.

La definizione di composizione di funzioni stabilisce che, date le funzioni f :

$A \rightarrow B$ e $g: B \rightarrow C$, la funzione composta $g \circ f$ fa corrispondere ad ogni elemento $a \in A$ l'elemento $g(f(a)) \in C$. Nel caso di funzioni parziali l'applicazione di tale definizione richiede cautela. Non sempre, infatti, due funzioni parziali $f: A \rightarrow B$ e $g: B \rightarrow C$ saranno tali che l'insieme delle immagini di f (che è un sottoinsieme di B) sia incluso nel dominio (di definizione) di g . In questo caso il dominio (di definizione) di $g \circ f$ potrebbe essere più piccolo dell'originale dominio (di definizione) di f . Ciò non accade per le funzioni (totali).

Esempio 3.9. Siano date le funzioni $f: \mathbb{R} \rightarrow \mathbb{R}$ e $g: \mathbb{R} \rightarrow \mathbb{R}$ tali che $f(x) = x^2$ e $g(x) = \sqrt{x-1}$. Chiaramente la f è una funzione totale, mentre il dominio di definizione della g è $Dg = \{x \in \mathbb{R} : x \geq 1\}$. Per definire la funzione composta $g \circ f$, teniamo presente che l'insieme delle immagini di f ,

cioè $\{x \in \mathbb{R} : x \geq 0\}$, non è un sottoinsieme del dominio di definizione di g . In questo caso dovremo escludere dal dominio di definizione della composta quei numeri x per i quali $f(x) \notin D_g$. Risulterà quindi $D_{go}f = \{x \in \mathbb{R} : x \geq 1\}$.

Questo fatto potrebbe portare alla conseguenza che la composta di due funzioni parziali abbia dominio di definizione vuoto, cioè risulti mai definita, come nel seguente esempio. Per quanto strano, ciò è perfettamente lecito.

Esempio 3.10. Siano date le funzioni f : $D \rightarrow \mathbb{R}$ (dove $D \subseteq \mathbb{R}$) e $g: \mathbb{R} \rightarrow \mathbb{R}$ tali che $f(x) = \sqrt{x}$ e $g(x) = -1 - x^2$. Per determinare la funzione composta $f \circ g$, osserviamo che il dominio di definizione della funzione f (che è la funzione che, essendo scritta per prima in $f \circ g$, opera per seconda) è $D_f = \{x \in \mathbb{R} : x \geq 0\}$. Notiamo

inoltre che l'insieme delle immagini della g (che opera per prima) è $Im(g) = \{x \in \mathbb{R} : x \leq -1\}$, in quanto, per ogni $x \in \mathbb{R}$, risulta $-1 - x^2 \leq -1$. Quindi, l'insieme delle immagini della funzione che opera per prima è costituito dai reali non maggiori di -1 , il dominio della funzione che opera per seconda è costituito dai reali non negativi, e tali insiemi sono disgiunti. Pertanto, nessun elemento del dominio della g avrà immagine appartenente al dominio della f : quindi nessun elemento avrà immagine nella funzione composta $f \circ g$. Quanto affermato può essere confermato dal ricavo (inutile) dell'espressione analitica di $f \circ g$: infatti, $(f \circ g)(x) = \sqrt{-1 - x^2}$. Concludendo, il dominio di $f \circ g$ nell'ambito dei numeri reali è \emptyset .

3.5 La funzione identità e la funzione inversa

Definizione 3.7. *Dato un insieme A , si dice identità su A e si indica con i_A (o semplicemente con i , quando A è chiaro dal contesto) la relazione di $A \times A$ che ad ogni elemento $a \in A$ associa se stesso, cioè $i_A = \{(a, a) : a \in A\}$.*

Lasciamo al lettore il compito di verificare che la i_A così introdotta è una funzione biiettiva, qualunque sia A . Inoltre, considerata una qualsiasi funzione $f: A \rightarrow B$, si verifica che: $f \circ i_A = i_B \circ f = f$.

Spesso è richiesto di determinare la relazione inversa di una funzione espressa nella forma $x \rightarrow f(x)$. Nell'esempio seguente è illustrato il procedimento per ricavare l'espressione della relazione inversa di una funzione data.

Esempio 3.11. *Si consideri la funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ tale che $f(x) = 2x$. Indichiamo con y il valore dell'espressione 2x , cioè*

l’immagine dell’elemento x . Essendo una relazione, la f è l’insieme di tutte e sole le coppie $(x, y) \in f$. Per ricavare la cercata espressione della relazione inversa, in base alla definizione, dobbiamo ottenere un insieme di coppie $(y, x) \in f^{-1}$, ovvero dobbiamo “scambiare” le posizioni ed i ruoli di x e y . Nel passaggio dalla funzione data alla sua relazione inversa, bisogna esplicitare dall’espressione $y = 2x$ la y , ottenendo $x = \frac{y}{2}$, che è l’immagine di x nella relazione inversa. Pertanto, possiamo affermare che la relazione inversa della funzione data è espressa dalla (nuova) funzione $f^{-1}(x) = \frac{x}{2}$.

Nel caso esaminato nell’esempio precedente, la relazione inversa della funzione f è un’altra funzione (cioè è una relazione tale da rispettare la definizione di funzione). Ma non sempre accadrà ciò, come mostrato nell’esempio seguente.

Esempio 3.12. Si consideri la funzione $g: \mathbb{R} \rightarrow \mathbb{R}$ tale che $g(x) = x^2$. Per ricavare la relazione inversa, la situazione è delicata: infatti la relazione inversa associa ad ogni elemento non negativo la sua radice quadrata, ma sia con il segno “+” che con il segno “−”. Insomma, essendo $g = \{(x, x^2): x \in \mathbb{R}\}$, si ha che $g^{-1} = \{(x^2, x): x \in \mathbb{R}\}$; da ciò risulta, per esempio, che $(9, 3) \in g^{-1}$ ma anche $(9, -3) \in g^{-1}$. Possiamo concludere che questa relazione inversa non rispetta la definizione di funzione. E ciò per ben due ragioni: primo, non tutti i numeri reali hanno una radice quadrata reale, ma soltanto i reali non negativi; in secondo luogo, per ogni reale positivo x , esistono due reali (e non uno solo) che elevati al quadrato danno come risultato x .

Visto che non sempre l'inversa di una funzione è una funzione, è interessante

studiare le funzioni che soddisfano tale proprietà.

Definizione 3.8. *Una funzione si dice invertibile funzione!invertibile quando anche la sua relazione inversa rispetta la definizione di funzione.*

Cerchiamo ora di scoprire quali caratteristiche deve avere una funzione affinchè anche la sua relazione inversa rispetti la definizione di funzione, cioè di caratterizzare le funzioni invertibili. A tale scopo, ricordiamo che, affinchè una relazione tra due insiemi sia una funzione, bisogna che ogni elemento del primo insieme (dominio) abbia una ed una sola immagine; ricordiamo inoltre che l'inversione di una relazione comporta lo scambio del ruolo del dominio del codominio. Quindi, per stabilire se la relazione inversa rispetta la definizione di

funzione dobbiamo esaminare se ogni elemento del dominio della relazione inversa

1. è dotato di almeno una immagine, e
2. è dotato di una sola immagine.

La prima condizione equivale a richiedere che ogni elemento del codominio della relazione originaria sia dotato di almeno una controimmagine; pertanto, la funzione originaria deve essere *suriettiva*. La seconda condizione equivale a richiedere che ogni elemento del codominio della relazione originaria sia dotato di una sola controimmagine; pertanto, la funzione originaria deve essere *iniettiva*. Possiamo concludere che le condizioni affinchè la relazione inversa di una funzione f rispetti la definizione di funzione sono l'iniettività e la suriettività della funzione originaria f : tale funzione,

quindi, deve essere *bijettiva*. Abbiamo quindi dimostrato la seguente

Proposizione 3.2. *Una funzione è invertibile se e soltanto se è bijettiva.*

Dunque una funzione bijettiva $f: D \rightarrow C$ associa ad ogni elemento di A uno ed un solo elemento di B e viceversa. Inoltre, se $f: A \rightarrow B$ è invertibile, si ricava semplicemente che $f^{-1} \circ f = i_A$ e $f \circ f^{-1} = i_B$.

Esercizi svolti

Esercizio 3.1. Sia $A = \{1, 2, 3\}$; si dica se le seguenti relazioni su A sono funzioni:

$$R_1 = \{(1, 2), (3, 1)\}$$

$$R_2 = \{(1, 2), (2, 3), (3, 1), (3, 3)\}$$

$$R_3 = \{(1, 2), (2, 3), (3, 1)\}$$

Soluzione: Una funzione da X in Y è una relazione $R \subseteq X \times Y$ tale che, per ogni x in X , esiste un unico y in Y tale che (x, y) appartiene a R . Pertanto, R_1 e R_2 non sono funzioni: la prima non associa 2 a nessun elemento mentre la seconda associa a 3 sia 1 che 3. Invece, R_3 è una funzione.

Esercizio 3.2. Sia $f(x) = x^2$ una funzione definita sull'intervallo chiuso dei reali $[-2, 8]$. Calcolare $f(4), f(-3)$ e $f(t - 3)$, per un'opportuna scelta del parametro t .

Soluzione: $f(4)$ vale 16; f non è definita per -3 e quindi $f(-3)$ non assume nessun valore. Infine, $f(t - 3)$ vale $(t - 3)^2$, per un qualsiasi $1 \leq t \leq 11$, visto che deve essere $-2 \leq t - 3 \leq 8$.

Esercizio 3.3. Sia $g: \mathbb{Z} \rightarrow \mathbb{Z}$ definita dalla formula $g(x) = |x| + 1$. Trovare l'insieme delle immagini di g .

Soluzione: L'immagine di g è l'insieme \mathbb{N}_+ .

Esercizio 3.4. Si consideri la funzione $f(x) = x$, dove $x \geq 0$. Indicare se ciascuna delle seguenti funzioni è un'estensione di f :

- (a) $g_1(x) = x$, dove $x \geq -2$
- (b) $g_2(x) = |x|$, per tutti gli $x \in \mathbb{R}$
- (c) $id: \mathbb{R} \rightarrow \mathbb{R}$
- (d) $g_3(x) = \frac{x+|x|}{2}$
- (e) $g_4(x) = x$, dove $x \in [-1, 1]$

Soluzione: Una funzione f' estende f (o, analogamente, f è una restrizione di f') se $dom(f) \subseteq dom(f')$ e per ogni $x \in dom(f)$ si ha che $f'(x) = f(x)$. Pertanto, g_1 , g_2 , id e g_3 estendono f , mentre g_4 non è un'estensione di f , visto che $dom(f) \not\subseteq dom(g_4)$.

Esercizio 3.5. Siano $A = [-3, 3]$, $B = [0, 3]$ e $C = [-3, 0]$; siano inoltre $f_1: A \rightarrow \mathbb{R}$, $f_2: B \rightarrow \mathbb{R}$ e $f_3: C \rightarrow \mathbb{R}$ definite dalla legge: “associa a ciascun numero il suo quadrato”. Quale delle funzioni date è iniettiva?

Soluzione: Sono iniettive soltanto f_2 ed f_3 : infatti, comunque presi x e y in B o in C si ha che, se $x \neq y$, allora $f(x) = x^2 \neq y^2 = f(y)$. Invece, comunque preso $x \in A \setminus \{0\}$, si ha che $f(x) = x^2 = f(-x)$, pur essendo $x \neq -x$.

Esercizio 3.6. Siano $A = [-3, 3]$, $B = [0, 9]$ e $C = [-27, 27]$; siano inoltre $f_1: A \rightarrow \mathbb{R}$, $f_2: A \rightarrow B$ e $f_3: A \rightarrow C$ definite dalla legge: “associa a ciascun numero il suo quadrato”. Quale delle funzioni date è suriettiva? Come cambierebbero le risposte se le funzioni associassero ad

ogni numero il loro cubo?

Soluzione: È suriettiva soltanto f_2 : infatti, l'immagine di f_i è B , per ogni $i \in \{1, 2, 3\}$. Se invece le f_i associano ad ogni numero il loro cubo, allora l'immagine di f_i è C ; quindi, soltanto f_3 è suriettiva.

Esercizio 3.7. *Sia $a \in \mathbb{R}$ e sia $D = \{x \in \mathbb{R} : x^2 \leq 2a^2 - 8\}$. Per quali valori di a la funzione $f: D \rightarrow \mathbb{R}$ definita da $f(x) = 5$ è iniettiva? È suriettiva?*

Soluzione: Essendo f una costante, l'insieme immagine di f è $\{5\}$ e quindi f non è suriettiva, qualunque sia a . Affinchè f sia iniettiva, D deve contenere al più un elemento; questo si ha soltanto quando $2a^2 - 8 \leq 0$, cioè per ogni $a \in [-2, 2]$.

Esercizio 3.8. *Siano le funzioni f e g definite sui numeri reali da $f(x) = x^2 + 2x$*

-3 e $g(x) = 3x - 4$. Trovare le formule che definiscono le funzioni composte $g \circ f$ e $f \circ g$.

Soluzione: Per definizione, $(g \circ f)(x) = g(f(x)) = g(x^2 + 2x - 3) = 3(x^2 + 2x - 3) - 4 = 3x^2 + 6x - 13$ e $(f \circ g)(x) = f(g(x)) = f(3x - 4) = (3x - 4)^2 + 2(3x - 4) - 3 = 9x^2 - 18x + 5$.

Esercizio 3.9. Siano $f(x) = k + 5 + x^2$ e $g(x) = \sqrt{x}$, dove \sqrt{x} è la funzione che restituisce la radice quadrata del numero x . Per quali valori di k sia $g \circ f$ che $f \circ g$ sono definite?

Soluzione: Anzitutto, osserviamo che $dom(g)$ è l'insieme dei reali non negativi. Pertanto, affinchè $g \circ f$ sia definita, deve essere che $k + 5 + x^2 \geq 0$; quindi, $k \geq -5$. Invece, $dom(f) = \mathbb{R}$ e quindi, qualunque sia k , $f \circ g$ è definita.

Esercizio 3.10. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ definita da $f(x) = 2x - 3$. Si dimostri che f è biiettiva, per cui ha una funzione inversa $f^{-1}: \mathbb{R} \rightarrow \mathbb{R}$; trovare una formula che definisca f^{-1} .

Soluzione: Banalmente, f è iniettiva: infatti, ogni volta che $f(x) = f(y)$ deve essere $x = y$, poichè $2x - 3 = 2y - 3$ implica che $x = y$. Inoltre, f è suriettiva; infatti, comunque fissiamo un reale y , possiamo sempre trovare un x tale che $y = f(x) = 2x - 3$: basta prendere $x = \frac{y+3}{2}$. Da ciò otteniamo che $f^{-1}(y) = \frac{y+3}{2}$.

Esercizio 3.11. Sia $f: \mathbb{Z} \rightarrow \mathbb{Z}$ definita da $f(x) = (-1)^x$. Si dimostri che f non è biiettiva; si definisca poi una opportuna restrizione di f che risulti invece biiettiva e si definisca per tale funzione la sua inversa. Quante restrizioni di f sono invertibili?

Soluzione: La funzione data non è iniettiva, visto che $f(2) = 1 = f(4)$, e non è suriettiva, visto che la sua immagine è $\{-1, 1\}$. Una possibile restrizione biiettiva di f è $g: \{0, 1\} \rightarrow \{-1, 1\}$ e la sua inversa è la funzione che mappa 1 in 1 e 0 in -1. Esistono infinite restrizioni di f invertibili, in particolare tutte le funzioni del tipo $\{a, b\} \rightarrow \{-1, 1\}$ tale che a è pari e b è dispari.

Esercizi da svolgere

Esercizio 3.12. Sia la funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ definita da

$$f(x) = \begin{cases} 3x - 1 & \text{se } x > 3 \\ x^2 - 2 & \text{se } -2 \leq x \leq 3 \\ 2x + 3 & \text{se } x < -2 \end{cases}$$

Calcolare $f(2), f(4), f(-1)$ e $f(-3)$.

Esercizio 3.13. Sia $W = \{1, 2, 3, 4\}$ e sia f una funzione da W a W definita da $f(1) =$

$1, f(2) = 3, f(3) = 1$ e $f(4) = 1$. Trovare l'insieme delle immagini di f .

Esercizio 3.14. Sia $a \in \mathbb{R}$. Per quali valori di a la funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ definita da $f(x) = a(x + 1)$ è iniettiva? È suriettiva?

Esercizio 3.15. Siano $A = \{a_1, \dots, a_m\}$ e $B = \{b_1, \dots, b_n\}$; sia inoltre f una funzione da A a B . Si dimostri che: (i) se f è iniettiva, allora $m \leq n$; (ii) se f è suriettiva, allora $m \geq n$; (iii) se f è biiettiva, allora $m = n$.

Esercizio 3.16. Siano $A = \{a_1, \dots, a_m\}$ e $B = \{b_1, \dots, b_n\}$; allora esiste una funzione biiettiva da A a B .

Esercizio 3.17. In termini insiemistici, qual'è la relazione fra una funzione $f: A \rightarrow B$ e la restrizione f di f ad un sottoinsieme A di A ?

Esercizio 3.18. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ definita da $f(x) = x^3$. Si dimostri che f è biiettiva e si determinino le funzioni $f \circ f^{-1}$, $f \circ f \circ f^{-1}$ o $f^{-1}, f \circ f \circ f \circ f^{-1}$ o $f \circ f \circ f$.

Esercizio 3.19. Si dimostri che una funzione $f: A \rightarrow B$ è biiettiva se e soltanto se esiste una funzione $g: B \rightarrow A$ tale che $g \circ f = i_A$ e $f \circ g = i_B$.

Esercizio 3.20. Si dimostri per esercizio la Proposizione 3.1.

Cardinalità

4.1 Equipotenza tra insiemi

Definizione 4.1. Due insiemi A e B si dicono equipotenti se sono in corrispondenza biunivoca, cioè quando esiste una funzione biiettiva di dominio A ed il codominio B .

Esempio 4.1. Considerato un triangolo qualsiasi, l'insieme A dei suoi lati e l'insieme B delle sue mediane sono equipotenti. Per verificare tale affermazione è sufficiente considerare la

relazione $R \subseteq A \times B$ che ad ogni lato $a \in A$ fa corrispondere la mediana $b \in B$ avente per estremi il vertice opposto al lato a ed il punto medio dello stesso lato a . La relazione R è una funzione: infatti, ad ogni lato del triangolo corrisponde, nella relazione introdotta, una ed una sola mediana. Questa funzione è inoltre suriettiva, giacchè ogni mediana è corrispondente, nella R , di un lato (quello avente come punto medio l'estremo della mediana non coincidente con un vertice del triangolo); è inoltre iniettiva, in quanto a due lati distinti corrispondono due distinte mediane (i loro estremi non coincidenti con un vertice sono i punti medi di due lati distinti): R è una funzione biiettiva. In base alla definizione concludiamo che A e B sono equipotenti.

Esempio 4.2. *L'insieme $I = \{0, 1\}$ e l'insieme $J = \{10, 11, 12\}$ non sono*

equipotenti. Se consideriamo, ad esempio, la funzione $f: I \rightarrow J$ tale che $f(x) = x + 10$ ci rendiamo conto che è una funzione iniettiva ma non suriettiva (il suo insieme delle immagini è $\{10, 11\}$ e non coincide con l'intero J), quindi non biiettiva. Non esiste alcuna funzione biiettiva $I \rightarrow J$, poichè ogni funzione da I a J non può essere suriettiva (altrimenti dovrebbe associare due immagini allo stesso elemento del dominio e non sarebbe, quindi, una funzione). Gli insiemi I e J non sono dunque equipotenti.

Come si può notare, la nozione di equipotenza esprime in maniera rigorosa quanto intendiamo intuitivamente dicendo: “due insiemi hanno lo stesso numero di elementi”. Una simile espressione presupporrebbe di aver già definito cosa sia un numero e varrebbe soltanto per insiemi finiti. Noi, invece

vogliamo procedere in generale.

Nel seguito considereremo gli insiemi dei quali ci siamo finora occupati (ed altri insiemi che possiamo immaginare) come elementi di un insieme universo U .

Importanti motivazioni ci impediscono però assolutamente di parlare di “insieme costituito da *tutti* gli insiemi”: dunque, l’insieme U dovrà essere considerato soltanto un insieme al quale appartengono, come elementi, altri insiemi, senza alcuna pretesa di totalità (v. [Esempio 4.5](#)).

Definiamo in U la relazione di equipotenza, precedentemente introdotta. Verificheremo che essa è una relazione di equivalenza, cioè che è riflessiva, simmetrica e transitiva.

Proposizione 4.1. *La relazione di equipotenza tra insiemi è una relazione di*

equivalenza.

Dimostrazione. Verifichiamo direttamente le tre proprietà della relazione di equivalenza.

1. La relazione di equipotenza tra insiemi è *riflessiva*. Infatti, ogni insieme $I \in U$ è equipotente a se stesso: si consideri, a tale riguardo, l'identità, che è una funzione biiettiva. Conseguentemente, in base alla Definizione 4.1, ogni insieme I risulta equipotente a se stesso.
2. La relazione di equipotenza tra insiemi è *simmetrica*. Infatti, se $I \in U$ è equipotente a $J \in U$, significa (per definizione) che esiste una funzione biiettiva $f: I \rightarrow J$; essendo biiettiva, f risulta invertibile, e la funzione inversa è $f^{-1}: J \rightarrow I$, anch'essa biiettiva (ad ogni elemento di J fa corrispondere uno ed un solo elemento di I e viceversa).

Pertanto, anche J è equipotente ad I .

3. La relazione di equipotenza tra insiemi è *transitiva*. Infatti, se $I \in U$ è equipotente a $J \in U$ ed inoltre $J \in U$ è equipotente a $L \in U$, significa (per definizione) che esistono due funzioni biettive $f: I \rightarrow J$ e $g: J \rightarrow L$.

Consideriamo la funzione composta $g \circ f: I \rightarrow L$: anch'essa è biettiva (ad ogni elemento di I fa corrispondere uno ed un solo elemento di L e viceversa).

Pertanto, I è equipotente a L .

□

La nozione di equipotenza tra insiemi ci permette innanzitutto di formulare quello che appare come un paradosso, ma è una proprietà che ci permette facilmente e rigorosamente di distinguere tra insiemi finiti ed infiniti. Sembra infatti molto strano che un insieme possa essere equipotente ad un suo sottoinsieme

proprio. Il lettore può rendersi conto di ciò verificando direttamente tale impossibilità nel caso di un insieme finito: ad esempio, non si può definire alcuna funzione biiettiva $f: A \rightarrow B$, essendo $A = \{0, 1\}$ e B il suo sottoinsieme proprio $\{1\}$.

Questo comportamento, però, non è caratteristico di *tutti* gli insiemi: esistono infatti insiemi che sono equipotenti ad alcuni loro sottoinsiemi propri.

Esempio 4.3. Consideriamo l'insieme $P = \{0, 2, 4, 6, 8, \dots\}$ dei numeri naturali pari (cioè, divisibili per 2) ed il suo sottoinsieme $I = \{2, 4, 6, 8, \dots\}$ dei numeri naturali pari e diversi da zero.

Mostreremo che, sebbene I sia un sottoinsieme proprio di P , i due insiemi P e I sono equipotenti. Consideriamo infatti la funzione $f: P \rightarrow I$ tale che $f(x) = x + 2$. Tale funzione è biiettiva: ad ogni $x \in P$ corrisponde, infatti, attraverso la

*funzione f , uno ed un solo $x + 2 \in I$;
viceversa, ad ogni elemento di I
corrisponde uno ed un solo elemento di P
attraverso la funzione inversa, cioè $f^{-1}(x)$
 $= x - 2$. Pertanto P ed I sono equipotenti.
Sottolineiamo ancora che I è un
sottoinsieme proprio di P : infatti è $P \setminus I =$
 $\{0\} \neq \emptyset$.*

Le definizioni di insieme finito e di insieme infinito si baseranno proprio sui due diversi comportamenti ora illustrati: in particolare, diremo *finito* ogni insieme che si comporterà analogamente a $\{0, 1\}$, diremo *infinito* ogni insieme che si comporterà analogamente a P . Possiamo dunque formalizzare quanto sopra introdotto nella seguente definizione.

Definizione 4.2. *Un insieme si dice infinito se è equipotente ad un suo sottoinsieme proprio. Un insieme si dice*

finito se non è equipotente ad alcun suo sottoinsieme proprio.

Ripercorreremo adesso la costruzione dell'insieme quoziante con riferimento all'equipotenza di insiemi. Concentriamo la nostra attenzione, in un primo momento, sugli insiemi *finiti*. Se indichiamo, come già sopra fatto, con U una famiglia avente per elementi insiemi (finiti ed infiniti), ma che contenga sicuramente tutti gli insiemi finiti, possiamo indicarla con $U_{fin} \subseteq U$ la sottofamiglia avente per elementi (soltanto) insiemi *finiti*. In U_{fin} consideriamo la relazione di equipotenza tra insiemi: la Proposizione 4.1 (formulata nel caso di insiemi qualsiasi) assicura che tale relazione è una relazione di equivalenza.

Consideriamo allora le classi di

equivalenza determinate in U_{fin} dalla relazione di equipotenza e consideriamo, infine, l'insieme quoziente avente tali classi di equivalenza come elementi. Non è difficile identificare ciascuna di tali classi di equivalenza con un numero naturale n (intuitivamente, il numero degli elementi appartenenti a ciascuno degli insiemi della classe di equivalenza); si dice che ogni elemento appartenente alla classe di equivalenza in questione ha *potenza* (o *cardinalità*) n . L'insieme quoziente può dunque essere interpretato come l'insieme \mathbb{N} dei numeri naturali intesi nel senso di numeri cardinali. Il naturale “zero” è dunque identificabile con la classe di equivalenza avente quale unico elemento l'insieme vuoto; potremo dire che l'insieme vuoto ha potenza 0.

Esempio 4.4. *Il numero naturale 2 si identifica con la classe di equivalenza*

(riferita alla relazione di equipotenza tra insiemi) alla quale appartiene l'insieme $I = \{0, 1\}$ e pertanto anche tutti gli insiemi ad esso equipotenti (l'insieme delle lenti di un comune paio di occhiali, l'insieme delle ruote di una bicicletta etc.). Ogni insieme appartenente alla classe di equivalenza sopra indicata ha potenza 2.

Indichiamo con $\# I$ il cardinale di un insieme I che identificherà la famiglia degli insiemi J equipotenti ad I . Ad esempio, se

$$A = \{3, 7\} \quad B = \{\alpha, \beta\} \quad C = \{\heartsuit, \spadesuit\}$$

abbiamo che $\# A = \# B = \# C$; spesso scriveremo anche $A = \# B = \# C$. Grazie alla Proposizione 4.1, abbiamo che $=_{\#}$ è una relazione d'equivalenza. Possiamo introdurre anche una relazione d'ordine tra numeri cardinali dicendo che $\# D \leq \# E$ significa che D è equipotente ad un

sottoinsieme di E ; spesso scriveremo anche $D \leq \# E$. È possibile dimostrare che $\leq \#$ è effettivamente un ordine (largo): è un facile esercizio per quanto riguarda la proprietà riflessiva e quella transitiva, mentre non è affatto elementare per la proprietà antisimmetrica, che noi dimostreremo nell'[Esercizio 4.2](#) sotto opportune ipotesi semplificative.

4.2 La potenza del numerabile

Il concetto di equipotenza ha consentito di formulare una definizione di insieme infinito. Finora, però, non sono stati trattati gli insiemi infiniti: abbiamo introdotto \mathbb{N} come insieme delle classi di equivalenza determinate, in U_{fin} (l'insieme costituito da tutti gli insiemi finiti), dalla relazione di equipotenza. Possiamo convincerci anche facilmente che \mathbb{N} stesso

non è finito. Ci occuperemo allora specificamente degli insiemi *infiniti*: dato che la relazione di equipotenza è definita anche per questi insiemi, affronteremo il problema di confrontare cardinalità di insiemi infiniti.

Per prima cosa consideremo i ben noti sistemi numerici. A tale riguardo, sorge spontanea la domanda: possiamo affermare che tutti gli insiemi infiniti hanno la *stessa* potenza? Sarebbe corretto, ad esempio, concludere che la potenza di insiemi come \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} è “infinito”, senza ulteriori specificazioni? Come vedremo, infatti, una delle massime conquiste della teoria degli insiemi di Georg Cantor (1845-1918) sarà proprio identificabile in questa inedita possibilità di “confrontare” (di “classificare”) gli insiemi infiniti.

Mostriamo intanto che l’insieme \mathbb{N} dei

numeri naturali è un insieme infinito. Per fare ciò, proveremo che \mathbb{N} è equipotente ad un suo sottoinsieme proprio.

Proposizione 4.2. *L'insieme $P \subseteq \mathbb{N}$ dei numeri naturali pari è equipotente a \mathbb{N} .*

Dimostrazione. Sia P l'insieme dei numeri naturali pari, cioè $P = \{m \in \mathbb{N}: m = 2 \cdot n, \text{ per qualche } n \in \mathbb{N}\}$. Per dimostrare che P , sottoinsieme proprio di \mathbb{N} , è equipotente a \mathbb{N} , è necessario individuare una funzione biettiva $f: P \rightarrow \mathbb{N}$. Una possibile funzione è regolata dalla legge $f(x) = \frac{x}{2}$. Mostriamo ora che f è biettiva (ovvero che P e \mathbb{N} sono posti, attraverso f , in corrispondenza biunivoca). Rappresentiamo f nel modo seguente:

$$\begin{array}{ccccccccc}
 & 0 & 2 & 4 & 6 & 8 & 10 & \cdots & (P) \\
 f & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & & \\
 & 0 & 1 & 2 & 3 & 4 & 5 & \cdots & (\mathbb{N})
 \end{array}$$

È evidente che ad ogni elemento di P

corrisponde dunque uno ed un solo elemento di \mathbb{N} e viceversa; ciò prova che f è biiettiva (alternativamente, usando il risultato mostrato nell'Esercizio 3.19, è possibile mostrare la biiettività di f osservando che la funzione $g(x) = 2x$ è la sua inversa).

□

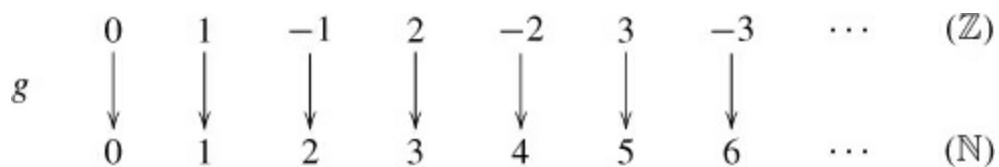
Possiamo ora introdurre una specifica denominazione per la potenza di tutti quegli insiemi che (come P) sono equipotenti a \mathbb{N} .

Definizione 4.3. *Un insieme equipotente a \mathbb{N} si dice avere la potenza del numerabile.*

L'insieme $P \subseteq \mathbb{N}$ dei naturali pari ha dunque la potenza del numerabile. La prossima proposizione darà un importante esempio di insieme equipotente a \mathbb{N} .

Proposizione 4.3. L'insieme \mathbb{Z} ha la potenza del numerabile.

Dimostrazione. Dobbiamo dimostrare che \mathbb{Z} è equipotente a \mathbb{N} , ovvero individuare una funzione biiettiva $g: \mathbb{Z} \rightarrow \mathbb{N}$. Tale funzione è così rappresentata schematicamente:



Dunque, g associa 0 a se stesso, gli interi positivi ai naturali dispari e gli interi negativi ai naturali pari maggiori di 0. Pertanto, ad ogni elemento di \mathbb{Z} corrisponde uno ed un solo elemento di \mathbb{N} e viceversa.

In forma analitica, g è definita da:

$$g(z) = \begin{cases} 0 & \text{se } z = 0 \\ 2(z - 1) + 1 & \text{se } z > 0 \\ -2z & \text{altrimenti} \end{cases}$$

Inoltre, si noti che non è difficile definire $g^{-1}: \mathbb{N} \rightarrow \mathbb{Z}$. Essa rispetta la legge:

$$g^{-1}(n) = \begin{cases} 0 & \text{se } n = 0 \\ \frac{n-1}{2} + 1 & \text{se } n \text{ è dispari} \\ -\frac{n}{2} & \text{altrimenti} \end{cases}$$

□

Quanto stabilito da queste due proposizioni potrebbe apparire sorprendente: un'interpretazione eccessivamente libera ed “intuitiva” del problema potrebbe infatti suggerire conclusioni ben diverse sulla “quantità” di elementi appartenenti agli insiemi esaminati. Ad esempio, l’insieme dei numeri naturali pari potrebbe sembrare costituito dalla “metà” degli elementi appartenenti all’insieme \mathbb{Z} ; analogamente, l’insieme \mathbb{Z} potrebbe sembrare costituito dal “doppio” degli elementi appartenenti a \mathbb{N} . Invece queste ultime “constatazioni”,

nonostante la loro apparente plausibilità, non hanno alcun significato in matematica: i tre insiemi \mathbb{N} , l'insieme P dei naturali pari e \mathbb{Z} sono equipotenti.

Occupiamoci ora dell'insieme \mathbb{Q} dei razionali. Intuitivamente, se confrontato con l'insieme dei numeri interi, \mathbb{Q} appare costituito da “moltissimi” elementi; potrebbe sembrare, a prima vista, che i razionali siano “molto più numerosi” dei naturali (e degli interi). Vale infatti la seguente

Proposizione 4.4. *Tra due naturali esistono infiniti razionali.*

Dimostrazione. Iniziamo con il mostrare che, dati due razionali a e b tali che $a < b$, esiste sempre un razionale q tale che $a < q < b$. Dall'assunzione $a < b$ possiamo derivare che $a+a < a+b$ e $a+b < b+b$ (sommendo membro a membro una volta a

e una volta b); pertanto, $2a < a + b < 2b$ e quindi $a < \frac{a+b}{2} < b$. La frazione $\frac{a+b}{2}$ il numero razionale q cercato. Siano ora n_1 e n_2 due numeri naturali diversi, con $n_1 < n_2$; essendo $\mathbb{N} \subset \mathbb{Q}$, possiamo considerare n_1 ed n_2 come numeri razionali. Per quanto dimostrato, esiste un $q \in \mathbb{Q}$ tale che $n_1 < q < n_2$. Possiamo ora iterare il ragionamento sulle due coppie di razionali (n_1, q) e (q, n_2) , ottenendo due razionali q_1 e q_2 tali che $n_1 < q_1 < q < q_2 < n_2$. Ora, iteriamo sulle coppie (n_1, q_1) , (q_1, q) , (q, q_2) e (q_2, n_2) ottenendo $n_1 < q < q_1 < q < q < q < q < q_2 < q < n_2$; e così via.

□

Eppure...

Proposizione 4.5. *L'insieme \mathbb{Q} ha la potenza del numerabile.*

Dimostrazione. Analogamente a quanto proposto nelle dimostrazioni precedenti, dobbiamo “allineare” in un’unica fila tutti gli elementi di \mathbb{Q} : fatto ciò, sarà possibile mettere in corrispondenza biunivoca gli elementi di \mathbb{Q} e quelli di \mathbb{N} . In realtà dimostreremo che $\mathbb{N} \times \mathbb{N}_0$ ha la potenza del numerabile. I razionali non negativi \mathbb{Q}_+ sono in corrispondenza biunivoca con un sottoinsieme di $\mathbb{N} \times \mathbb{N}_0$, si veda l’Esercizio 2.13. Se, dunque, dimostriamo che $\mathbb{N} \times \mathbb{N}_0$ ha la potenza del numerabile, \mathbb{Q}_+ dovrà avere la stessa potenza, dal momento che contiene sicuramente un sottoinsieme equipotente ad \mathbb{N} . Scriviamo gli elementi di $\mathbb{N} \times \mathbb{N}_0$ nella tabella seguente:

| | | | | | |
|---------------|---------------|---------------|---------------|---------------|----------|
| $\frac{0}{1}$ | $\frac{0}{2}$ | $\frac{0}{3}$ | $\frac{0}{4}$ | $\frac{0}{5}$ | \dots |
| $\frac{1}{1}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{1}{5}$ | \dots |
| $\frac{2}{1}$ | $\frac{2}{2}$ | $\frac{2}{3}$ | $\frac{2}{4}$ | $\frac{2}{5}$ | \dots |
| $\frac{3}{1}$ | $\frac{3}{2}$ | $\frac{3}{3}$ | $\frac{3}{4}$ | $\frac{3}{5}$ | \dots |
| $\frac{4}{1}$ | $\frac{4}{2}$ | $\frac{4}{3}$ | $\frac{4}{4}$ | $\frac{4}{5}$ | \dots |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |

Grazie a questa tabella, possiamo scrivere una “fila” di coppie:

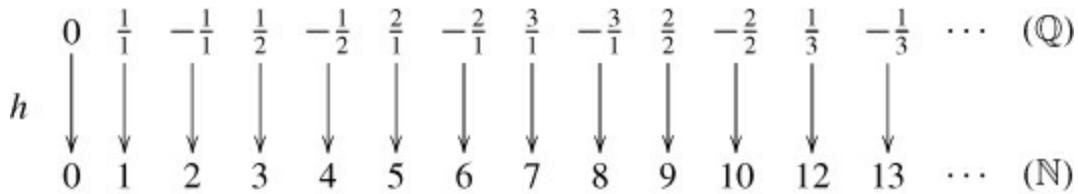
- partiamo dall’elemento $\frac{0}{1}$ (che si trova, nella tabella, in alto a sinistra);
- percorriamo la tabella procedendo “a zig-zag”, ovvero secondo una serpentina che, da $\frac{0}{1}$, individua successivamente:
 $\frac{0}{1}, \frac{0}{2}, \frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{0}{3}, \frac{0}{4}, \frac{1}{3}, \frac{2}{2}, \frac{3}{1}, \dots$ (provare per esercizio a scrivere la funzione da \mathbb{N} a \mathbb{Q}_+ che “enumeri” questi elementi).

Otteniamo, pertanto, quanto inizialmente ricercato: l’intero insieme $\mathbb{N} \times \mathbb{N}_0$ risulta ordinato nella “fila” precedente; pertanto,

\mathbb{Q}_+ sarà ordinabile eliminando dalla fila precedente le frazioni “apparenti”, tipo

$$\frac{2}{2}, \frac{3}{3}, \frac{2}{4}, \dots$$

Per mettere in corrispondenza biunivoca tutto q con \mathbb{N} mettiamo nella “fila” degli elementi di $\mathbb{N} \times \mathbb{N}_0$ dopo ciascuna frazione $\frac{m}{n}$ la frazione negativa $-\frac{m}{n}$. Otteniamo così la funzione biiettiva $h: \mathbb{Z} \times \mathbb{N}_0 \rightarrow \mathbb{N}$ rappresentata da:



Come prima, eliminando da questo ordinamento le frazioni apparenti, abbiamo una numerazione di \mathbb{Q} .

□

In conclusione, abbiamo verificato che tutti gli insiemi numerici finora considerati (l’insieme dei naturali pari, \mathbb{Z} e

\mathbb{Q}) hanno la potenza del numerabile.

4.3 La potenza del continuo

Quanto affermato nel paragrafo precedente pone una questione centrale: *tutti* gli insiemi infiniti hanno la potenza del numerabile? La risposta è: *no*. La proposizione seguente, dovuta a Cantor, motiva questa risposta.

Proposizione 4.6. *L'insieme \mathbb{R} non ha la potenza del numerabile.*

Dimostrazione. Faremo vedere che un sottoinsieme di \mathbb{R} ha una potenza superiore al numerabile. Consideriamo l'intervallo reale $(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}$; un numero r di tale insieme sarà della forma $r = 0, a_0 a_1 a_2 \dots a_n \dots$, dove ogni a_i è una cifra decimale (cioè, $a_i \in \{0, 1, \dots, 9\}$)

per ogni i), cioè $r = a_1 \cdot 10^{-1} + a_2 \cdot 10^{-2} + \dots + a_n \cdot 10^{-n} + \dots$. In modo simile, scriviamo in base 2 i numeri in esso contenuti; in questo modo ci accorgiamo che possiamo descriverli con successioni numerabili della forma $b_1 b_2 \dots b_n \dots$, con $b_i \in \{0, 1\}$, cioè le cifre che compaiono dopo la virgola nella rappresentazione binaria. Questa successione sarà definitivamente uguale a 0 se il numero è razionale, non lo sarà se il numero è irrazionale (per evitare ogni ambiguità, possiamo scegliere di non considerare successioni costituite, da un certo punto in poi, da una fila illimitata di sole cifre 1: com'è consuetudine, ad esempio, 0, 10111111 ... è considerato uguale a 0, 11000000 ...).

Ammettiamo, per assurdo, che l'insieme di tali successioni abbia la potenza del numerabile: dovrebbe allora essere possibile, analogamente a quanto fatto

nelle dimostrazioni precedenti, “allineare” in un’unica fila tutte queste successioni:

$$b_1^1 \ b_2^1 \ \cdots \ b_n^1 \ \cdots$$

$$b_1^2 \ b_2^2 \ \cdots \ b_n^2 \ \cdots$$

⋮

$$b_1^n \ b_2^n \ \cdots \ b_n^n \ \cdots$$

⋮

Scegliamo ora la diagonale della tabella $b_1^1 b_2^2 \dots b_n^n \dots$: anche questa sarà una $1 \ 2 \ n$ successione numerabile di cifre binarie e così lo sarà anche quella ottenuta da essa scambiando ogni 0 con un 1 e viceversa. In questo modo avremo identificato un numero reale $r = \bar{b}_1^1 \bar{b}_2^2 \dots \bar{b}_n^n \dots$ (dove \bar{b} è il complemento del bit b) appartenente $1 \ 2 \ n$ a (0, 1) e sicuramente diverso da ognuno di quelli elencati precedentemente.

Abbiamo contraddetto l’affermazione che pretendeva di aver già elencati tutti i numeri reali compresi tra 0 e 1. Possiamo

pertanto concludere che è impossibile “elencare” in un’unica “fila” tutti i numeri reali compresi tra 0 e 1; quindi, $(0, 1)$ *non* può essere messo in corrispondenza biunivoca con \mathbb{N} e *non* ha la potenza del numerabile. Di conseguenza, è impossibile “elencare” in un’unica “fila” tutti gli elementi di \mathbb{R} .

Possiamo anche far vedere, con una costruzione geometrica, che \mathbb{R} e $(0, 1)$ sono equipotenti. Consideriamo la retta reale e la semicirconferenza inferiore C senza estremi, di raggio $\frac{1}{2}$ e di centro $(\frac{1}{2}, \frac{1}{2})$. È facile, con una proiezione centrale dal centro di C , stabilire una corrispondenza biunivoca tra i punti di C e la retta dei reali: al punto $c \in C$ facciamo corrispondere il reale r_c ottenuto intersecando la retta dei reali con la retta passante per c e per il centro di C . È altresì possibile, con una proiezione ortogonale, stabilire una corrispondenza

biunivoca tra C ed il segmento $(0, 1)$: al punto $c \in C$ facciamo corrispondere il reale $x_c \in (0, 1)$ ottenuto intersecando la retta dei reali con la retta verticale passante per c . La biezione cercata tra $(0, 1)$ e \mathbb{R} è quella che associa x_c con r_c .

□

Corollario 4.1. *L'insieme delle parti di \mathbb{N} è equipotente a \mathbb{R} .*

Dimostrazione. Basta osservare che un sottoinsieme (anche infinito) di naturali A può essere rappresentato da una sequenza infinita di bit $b_0 b_1 b_2 \dots b_n \dots$, dove $b_i = 1$ se e soltanto se $i \in A$; tale rappresentazione viene chiamata *rappresentazione caratteristica* dell'insieme A e la corrispondente funzione $\chi_A: \mathbb{N} \rightarrow \{0, 1\}$, tale che $\chi_A(i) = 1$ se e soltanto se $i \in A$, è chiamata *funzione caratteristica* di A . Pertanto, $\wp(\mathbb{N})$, l'insieme delle parti di \mathbb{N} , è equipotente

all’intervallo $(0, 1)$, che a sua volta è equipotente a \mathbb{R} .



La proposizione ora dimostrata richiede evidentemente che la potenza di \mathbb{R} venga denominata con un’espressione diversa da “potenza del numerabile”: la definizione seguente stabilisce tale nuova denominazione.

Definizione 4.4. *Un insieme equipotente a \mathbb{R} si dice avere la potenza del continuo.*

Abbiamo potuto constatare direttamente che non tutti gli insiemi infiniti hanno la stessa potenza, in particolare che non tutti gli insiemi infiniti sono equipotenti a \mathbb{N} : la potenza dell’insieme \mathbb{R} , cioè del continuo, è diversa dalla potenza del numerabile. A questo punto, è spontaneo porsi il problema: esistono insiemi infiniti non

equipotenti nè a \mathbb{N} nè a \mathbb{R} ?

La matematica contemporanea ha dato risposta affermativa a questa domanda: è possibile costruire un numero qualsiasi di altri insiemi infiniti aventi potenze sempre diverse (potremmo dire, intuitivamente, sempre più “numerosi”). Indichiamo con il simbolo \aleph_0 (che si legge: *Aleph con zero*) la potenza del numerabile e con $2\aleph_0$ la potenza del continuo: essi sono detti *numeri transfiniti*. È possibile definire infiniti numeri transfiniti iterando la costruzione dell’“insieme delle parti”; pertanto, l’insieme dei numeri transfiniti è, a sua volta, un insieme infinito.

Ma esistono numeri transfiniti compresi tra \aleph_0 e $2\aleph_0$? Esiste cioè un insieme infinito $I \subset \mathbb{R}$ che non abbia né la potenza del numerabile nè la potenza del continuo? Per molto tempo non si era

riusciti a dare una risposta a questa domanda; in particolare, era stata formulata la seguente congettura:

Ogni sottoinsieme infinito di \mathbb{R} non avente la potenza del numerabile ha la potenza del continuo.

Essa, detta *ipotesi del continuo*, equivale a negare l'esistenza di transfiniti intermedi tra \aleph_0 e $2\aleph_0$. Nel 1962, Paul Joseph Cohen (1934-2007) dimostrò che l'ipotesi del continuo appartiene ad una particolare classe di questioni, denominate *indecidibili*, nel senso che nè essa nè la sua negata sono dimostrabili nei classici sistemi assiomatici per la teoria degli insiemi e, se aggiunte (individualmente) come assiomi a tali sistemi, non portano a contraddizioni.

4.4 Cenni sulle antinomie della teoria degli insiemi

4.4.1 Le antinomie

Concludiamo presentando alcune importanti considerazioni storiche, dedicate ad uno degli argomenti più affascinanti della logica matematica: ci riferiamo alla questione delle *antinomie*, situazioni decisamente contraddittorie da distinguere dai paradossi, enunciati apparentemente ingannevoli, strani, ma eleganti, nei quali sembrano “convivere” verità e falsità (la letteratura su antinomie e paradossi è vastissima; ci limitiamo ad indicare [17, 23]). Paradossale è, ad esempio che un insieme sia equipotente ad una sua parte propria, ma, come abbiamo visto, ciò è rigorosamente dimostrabile.

Il termine *antinomia* deriva dal greco “anti” (contro) e “nomos” (legge, norma); esso è utilizzato per indicare il verificarsi di una contraddizione, tipo la presenza di una particolare proposizione che al contempo risulta essere vera e falsa nell’ambito di una teoria. L’improponibile contrasto tra i valori di verità risulta allora inevitabile, finendo così per ledere la stessa validità del sistema logico nel quale viene scoperta l’antinomia in questione.

Verso la fine del secolo XIX, vengono proposte ed esaminate numerose antinomie riguardanti la teoria degli insiemi. Alcune di esse sono antinomie semantiche, ovvero collegate al significato attribuito a particolari termini del linguaggio; altre invece sono antinomie logiche, ovvero connesse alla struttura delle frasi in questione, come la celebre antinomia di Russell, che esamineremo

nel paragrafo seguente.

“*L'enunciato che segue è falso.
L'enunciato che precede è vero.*”

Congiuntamente questi enunciati danno lo stesso risultato del paradosso di Epimenide. Eppure, separatamente, sono enunciati innocui e perfino potenzialmente utili”.

Douglas R. Hofstaedter

Infatti una delle più celebri antinomie semantiche è *l'antinomia del mentitore* (o *di Epimenide*), nota sino dall'antichità [28, 37, 39, 48]. Essa è sintetizzata nella frase:

Io sto mentendo

Riflettendo su questa affermazione, ci rendiamo conto che:

1. se chi pronuncia tale frase dice la verità, allora egli sta effettivamente mentendo, e questa è una contraddizione;
2. se chi pronuncia tale frase mente, allora egli non sta mentendo e, di conseguenza, sta dicendo la verità: anche questa è una contraddizione.

Illustriamo ora una semplice ed elegante antinomia della denotazione (detta di Berry-Russell, ma la proposta originale è di Richard). È noto che il numero delle sillabe dei nomi con i quali si indicano, nella lingua inglese, i numeri interi positivi tende a crescere al crescere del numero nominato. Pertanto, i nomi di alcuni numeri interi positivi devono essere costituiti da almeno diciannove sillabe; fra questi, ne esisterà uno più piccolo di tutti

gli altri. Chiameremo tale numero (che, in inglese, è 111777) “il più piccolo intero non definibile con meno di diciannove sillabe” che (in inglese) è proprio una “definizione” costituita da diciotto sillabe! La contraddizione appare dunque evidente: “il più piccolo intero non definibile con meno di diciannove sillabe” è stato definito (in inglese) con sole diciotto sillabe. In questo caso, però, le radici dell’antinomia vanno ricercate nel particolare tipo di denotazione che si usa per i termini in questione.

Esempio 4.5. Ritorniamo ora alla questione dell’insieme di tutti gli insiemi. Anche questa nozione risulta contraddittoria. Infatti, ricordando che l’insieme delle parti $\wp(I)$ è l’insieme costituito da tutti i sottoinsiemi di I , si può dimostrare che $I <_{\#} \wp(I)$, qualunque sia I (Teorema di Cantor); ciò generalizza il

risultato presentato nel Corollario 4.1. Immaginiamo ora di poter parlare dell'insieme totale (insieme di “tutti gli insiemi”) T ; risulterebbe che $\wp(T) \subseteq T$ (infatti T è l’insieme totale!) e quindi $\wp(T) \leq_{\#} T$, contro il Teorema di Cantor, che porta invece a $T <_{\#} \wp(T)$.

4.4.2 Gottlob Frege e Bertrand Russell

Il progetto fondazionale di Gottlob Frege (1848-1925), uno dei massimi logici dell’intera storia della cultura, può essere sintetizzato nel tentativo di ricondurre interamente la matematica alla logica (attraverso la teoria degli insiemi). Nel 1902, proprio alla vigilia della pubblicazione della seconda parte della grande opera logica fregeana (*Principi dell’Aritmetica derivati*

ideograficamente), il trentenne Bertrand Russell (1872-1970) rilevò una contraddizione nel capolavoro del logico tedesco: essa è riassunta nella celebre *antinomia di Russell* o *antinomia degli insiemi normali* [18].

La formulazione originale dell'antinomia di Russell è basata sulla definizione di insieme normale, che illustreremo. L'idea di insieme, come sappiamo, non è introdotta da una definizione, ma è un concetto primitivo; non ci sono particolari restrizioni al tipo di elementi che possono appartenere all'insieme dato. È quindi possibile (perchè no?) richiedere che un certo insieme appartenga a se stesso come elemento. Ecco dunque la definizione di Russell: un insieme I si definisce *normale* quando ha la proprietà di non contenere se stesso come elemento.

Esempio 4.6. L'insieme $I = \{x \in \mathbb{N} : x < 5\} = \{0, 1, 2, 3, 4\}$ è un insieme normale, in quanto non contiene I stesso come elemento: $I \notin I$. L'insieme T di tutti gli insiemi di cardinalità maggiore di 2 non è un insieme normale, in quanto contiene se stesso come elemento, avendo T più di due elementi.

Consideriamo ora l'insieme N avente per elementi tutti e soltanto gli insiemi normali: N contiene se stesso come elemento?

- Una prima possibilità è che N contenga se stesso come elemento. Ma N è l'insieme formato da tutti e soltanto gli insiemi normali, ovvero da tutti e soltanto quegli insiemi che non contengono se stessi come elemento; perciò dovremmo concludere che, se N appartiene a N , risulta che N non può

appartenere a se stesso come elemento!

La nostra prima risposta finisce quindi per essere contraddittoria.

- La rimanente possibilità è che N non appartenga a se stesso come elemento. Ma, in tale caso, N risulterebbe essere un insieme normale e, di conseguenza, dovrebbe proprio appartenere a N . Ed anche questa risposta è contraddittoria.

In simboli, l'insieme N è $\{I : I \notin I\}$ e le due possibilità sull'appartenenza dell'insieme N a stesso portano alle implicazioni:

se $\mathbb{N} \in N$ allora $\mathbb{N} \notin N$

se $\mathbb{N} \notin N$ allora $\mathbb{N} \in N$

ovvero a un'inevitabile contraddizione.

Nota 4.1. *Per molti versi analoga all'antinomia degli insiemi normali è*

l’antinomia del barbiere (lo stesso Bertrand Russell la ricorda come pressochè equivalente alla propria antinomia: si veda ad esempio [2]): “In un paese isolato, uno degli abitanti è il barbiere e rade tutti (e soltanto) coloro che non si radono da sè.” Domanda: in quel paese, chi rade il barbiere?

Ammettiamo che il barbiere si rada da sè; ma allora è proprio il barbiere che lo rade, mentre avevamo affermato che il barbiere rade tutti e soltanto coloro che non si radono da sè! Questa prima risposta è quindi contraddittoria.
Ammettiamo allora che il barbiere non si rada da sè; ma in tale caso dovrebbe essere proprio il barbiere a raderlo, in quanto avevamo affermato che il barbiere rade tutti e soltanto coloro che non si radono da sè: quindi egli si rade da sè. Ed anche questa seconda risposta si

rivela contraddittoria.

Che dire se il barbiere fosse una donna?

Esercizi svolti

In questa sezione assumeremo, per semplicità, di lavorare soltanto con insiemi finiti, numerabili o continui. Questa assunzione semplificativa ci permetterà di dimostrare formalmente molte proprietà che valgono per insiemi qualsiasi, ma la cui dimostrazione nel caso generale risulta troppo complicata per le finalità del presente testo.

Esercizio 4.1 (Teorema di Cantor-Bernstein). *Se esiste una iniezione da A a B ed una iniezione da B ad A, allora esiste una biiezione da A a B.*

Soluzione: Anzitutto osserviamo che A è finito se e soltanto se B è finito: infatti, non esistono funzioni iniettive da un insieme infinito ad uno finito. Sia quindi $\# A = n$ e $\# B = m$; per l'Esercizio 3.15, abbiamo che la funzione iniettiva da A a B implica che $n \leq m$, mentre la funzione iniettiva da B ad A implica che $m \leq n$. Quindi $m = n$ ed è facile costruire una biiezione da A a B .

Siano quindi A e B entrambi infiniti. Se sono entrambi numerabili o entrambi continui, è facile metterli in corrispondenza biunivoca: basta comporre le biiezioni tra essi e \mathbb{N} o \mathbb{R} , rispettivamente. Diciamo quindi per assurdo che A sia numerabile e B sia continuo (il viceversa è analogo); sia inoltre f la funzione iniettiva da B ad A . Visto che A è numerabile, esisterebbe una corrispondenza biunivoca tra B e un $N \subseteq$

\mathbb{N} (basta comporre f con la biiezione tra A e \mathbb{N} , che esiste essendo A numerabile). Ma è facile convincersi che N o è finito o è numerabile (la dimostrazione formale è lasciata al lettore nell'Esercizio 4.8); questo porta a contraddirre il fatto che B sia continuo.

Esercizio 4.2. *Si dimostri che $\leq_{\#}$ è una relazione d'ordine sugli insiemi.*

Soluzione:

- Banalmente, $\leq_{\#}$ gode della proprietà riflessiva, visto che $A \leq_{\#} A$, essendo l'identità una biiezione da A in se stesso.
- Per quanto riguarda la proprietà transitiva, supponiamo che $A \leq_{\#} B$ e $B \leq_{\#} C$; siano inoltre $f: A \rightarrow B'$ e $g: B \rightarrow C'$ le biiezioni tra A (risp., B) e un sottoinsieme B' (risp., C') di B (risp., di C). Consideriamo g' , la restrizione di g a

B , e la funzione composta $g' \circ f$; quest'ultima è una biiezione da A in C' , da cui $A \leq_{\#} C$.

- Per quanto riguarda la proprietà antisimmetrica, supponiamo che se $A \leq_{\#} B$ e $B \leq_{\#} A$. Questo implica che esiste una funzione iniettiva da A in B e da B in A : infatti ogni funzione biettiva da X in $Y' \subseteq Y$, vista come funzione da X in Y , è iniettiva. Per il Teorema di Cantor-Bernstein, esiste una biiezione tra A e B ; quindi, $A =_{\#} B$.

Esercizio 4.3. *Dimostrare che l'intervallo chiuso $[0, 1]$ è equipotente all'intervallo aperto $(0, 1)$.*

Soluzione: È facile dimostrare che $(0, 1) \leq_{\#} [0, 1]$, essendo $(0, 1) \subseteq [0, 1]$ (si veda l'Esercizio 4.9); per lo stesso motivo, $[0, 1] \leq_{\#} (-\infty, +\infty)$. Visto che nella Proposizione

4.6 abbiamo dimostrato che $(0, 1) =_{\#} (-\infty, +\infty)$, abbiamo che $(0, 1) \leq_{\#} [0, 1] \leq_{\#} (0, 1)$; essendo $\leq_{\#}$ una relazione d'ordine, la proprietà antisimmetrica implica che $(0, 1) =_{\#} [0, 1]$.

Esercizio 4.4. *Siano A e B due insiemi infiniti tali che $A =_{\#} B$; allora $A \cup B =_{\#} A (=_{\#} B)$.*

Soluzione: Siano A e B numerabili; esiste allora una biiezione f tra A ed \mathbb{N} e una biiezione g tra B ed \mathbb{N} . Come abbiamo visto nella Proposizione 4.2, esiste una biiezione p da \mathbb{N} ai naturali pari e, similmente alla Proposizione 4.2, si può dimostrare che esiste una biiezione q da \mathbb{N} ai naturali dispari. Consideriamo le funzioni $p \circ f$ e $q \circ g$; esse sono biiezioni (perchè composizione di biiezioni, si veda la Proposizione 3.1) che associano A ai

naturali pari e B ai dispari. Pertanto, la loro unione è ancora una biiezione da $A \cup B$ a \mathbb{N} .

Siano A e B continui; ripetiamo la stessa costruzione, ma considerando \mathbb{R}_+ , i reali non negativi, e \mathbb{R}_- , i reali negativi. Infatti, seguendo un procedimento simile all'esercizio precedente, abbiamo che $\mathbb{R}_+ =_{\#} \mathbb{R}_- =_{\#} \mathbb{R}$.

Esercizio 4.5. *Si dimostri che, se B non è finito, allora, comunque scelto $B' \subseteq B$, si ha che $B' =_{\#} B$ oppure che $B \setminus B' =_{\#} B$.*

Soluzione: Diciamo per assurdo che esista un $B' \subseteq B$ tale che $B' \neq_{\#} B$ e $B \setminus B' \neq_{\#} B$.

Poichè sia B' che $B \setminus B'$ sono sottoinsiemi di B , sia ha che $B' \leq_{\#} B$ e $B \setminus B' \leq_{\#} B$; deve pertanto essere che $B' <_{\#} B$ e $B \setminus B' <_{\#} B$.

Distinguiamo due casi:

- se B è numerabile, allora B' e $B \setminus B'$

devono essere finiti; allora la loro unione (cioè B) dovrebbe essere un insieme finito: contraddizione!

- se B è continuo, allora B' e $B \setminus B'$ saranno al più numerabili; allora la loro unione (cioè B) dovrebbe essere un insieme numerabile (si veda l'esercizio precedente): contraddizione!

Esercizi da svolgere

Esercizio 4.6. *La sequenza $\{(1, 1), (4, 8), (9, 27), \dots, (n^2, n^3), \dots\}$ è numerabile?*

Esercizio 4.7. *Gli intervalli $G = [0, 1]$ e $H = [2, 5]$ sono equipotenti?*

Esercizio 4.8. *Si dimostri che ogni sottoinsieme di \mathbb{N} o è finito o è numerabile (N.B.: questo esercizio può essere risolto facilmente anche senza ricorrere al*

Teorema di Cantor-Bernstein).

Esercizio 4.9. *Dimostrare che, se $A \subseteq B$, allora $A \leq_{\#} B$.*

Esercizio 4.10. *Sia $A \subseteq \mathbb{N}$; si dimostri che, se $\mathbb{N} \setminus A$ è finito, allora $A =_{\#} \mathbb{N}$. Cosa si può dire dell'implicazione inversa? Cosa cambierebbe se al posto di \mathbb{N} considerassimo \mathbb{R} ?*

Esercizio 4.11. *Siano A e B due insiemi finiti; si dimostri che:*

1. $\#(A \cup B) = \# A + \# B - \#(A \cap B)$;
2. $\#(A \times B) = \# A \cdot \# B$;
3. $\#(B^A) = \# B^{\# A}$, dove B^A è l'insieme di tutte le funzioni da A in B .

Esercizio 4.12. *Sia A un qualunque insieme e si consideri l'insieme di tutte le funzioni da A in $\{0, 1\}$ (tale insieme viene*

solitamente indicato con $\{0, 1\}^A$). Si dimostri che $\{0, 1\}^A$ è equipotente a $\wp(A)$, l'insieme delle parti di A .

Esercizio 4.13. *Si dimostri che, se A è finito, allora $\#(\wp(A)) = 2^{\#A}$.*

Esercizio 4.14. *Sia A un insieme finito e sia Bij_A l'insieme di tutte le biiezioni di A in sè. Allora, $\#(Bij_A) = (\# A)!$.*

5

Numeri naturali

5.1 L'insieme dei numeri naturali

Nel precedente capitolo abbiamo introdotto i singoli numeri naturali come classi di equipotenza di insiemi finiti, cioè come numeri cardinali, mettendo in risalto l'aspetto quantitativo. L'insieme dei numeri naturali è stato definito poi come la collezione di tali numeri.

In realtà non è questo il modo più “naturale” di introdurre i numeri naturali. Probabilmente l'intuizione di numero naturale è più legata al procedimento del

contare che al concetto di quantità. Già dall'antichità classica si parlava di *serie naturale dei numeri*. Comunque il concetto di numero ordinale è almeno altrettanto fondamentale di quello di numero cardinale. In particolare la struttura del sistema ordinato dei numeri si è rivelata cruciale sia nel mondo della matematica che in quello dell'informatica.

A differenza di quanto accaduto con la geometria, gli antichi matematici non si posero il problema di caratterizzare i numeri naturali, probabilmente perché per loro essi avevano uno status metafisico che li rendeva addirittura più reali degli oggetti fisici (si pensi a Pitagora e Platone).

Nel Medioevo essi perdono tale status e ci si rende conto che sono, o almeno li si tratta, come semplici termini di un linguaggio. Nasce allora l'esigenza di una loro caratterizzazione e si tenta una prima

assiomatizzazione. Non soltanto come curiosità storica vale la pena di dare uno sguardo al primo tentativo che ci è stato tramandato in questo senso. Esso è dovuto a Campano da Novara (seconda metà del XIII secolo), il quale, nella sua traduzione/edizione degli *Elementi* di Euclide, formulò la seguente definizione [Labella, 2000]: “*Si dice serie naturale dei numeri quella per la quale il calcolo degli stessi avviene aggiungendo un’unità.*” A questa definizione, egli aggiungeva i seguenti assiomi (*petitiones*):

1. Di ogni numero si possono prendere quante copie o quanti multipli si vuole.
2. La serie dei numeri può procedere all’infinito.
3. Nessun numero può essere diminuito all’infinito.

Si osservi che il primo assioma ci

permette di affermare che ormai il numero è trattato come un termine del linguaggio (e quindi può essere ripetuto quante volte si vuole). I successivi due assiomi, anche se assolutamente insufficienti a caratterizzare il sistema dei numeri naturali con il rigore richiesto dall'attuale pratica matematica, ne esprimono però molto efficacemente la struttura. Il secondo ne caratterizza il procedimento costruttivo di generazione, mentre l'ultimo è una forma della moderna *proprietà del buon ordinamento* che può essere una forma alternativa del *principio di induzione*, come vedremo più avanti. Naturalmente manca la richiesta (per quell'epoca inconcepibile) che nel diminuire un numero ci si fermi sempre allo stesso elemento, cioè sia unico il punto di partenza. Questa “assiomatizzazione” fu presto dimenticata quando, con l'avvento del Rinascimento, si

cercò di eliminare tutto ciò che sapeva di medioevale. Tuttavia la concezione di numero naturale come termine del linguaggio era ormai acquisita.

Alla fine del XIX secolo, l'intero mondo matematico è impegnato nella puntualizzazione rigorosa dei fondamenti della disciplina: le teorie di George Boole (1815-1864) e di Georg Cantor (1845-1918) e le approfondite ricerche logicomatematiche di Gottlob Frege sono testimonianze chiare dello spirito che pervade un ampio settore della ricerca matematica di quel periodo.

Ci siamo già occupati dell'opera di Frege nella presentazione dell'antinomia di Russell; nel 1884, Frege pubblica *Die Grundlagen der Arithmetik*, l'opera in cui si trova la fondamentale definizione di numero. Egli introduce il numero “zero” facendolo corrispondere al concetto di

“diverso da se stesso” (ovvero ad una condizione di impossibilità): potremmo dire che, per Frege, “zero” è la quantità di elementi che verificano una condizione impossibile.

A partire dallo zero, Frege introduce ricorsivamente ogni altro naturale, ciascuno sulla base del precedente: così il numero 1 è associato al (solo) concetto di “zero” (si tratta dunque di *un* concetto, considerato singolarmente), il numero 2 ai (*due*) concetti di “zero” e di “uno”, il numero 3 ai (*tre*) concetti di “zero”, di “uno” e di “due”, e così via.

Ricordiamo che, sette anni dopo la pubblicazione del trattato di Frege, un grande matematico italiano propone per l’aritmetica una sistemazione di tipo assiomatico. Con il lavoro *Sul concetto di numero*, del 1891, Giuseppe Peano (1858-1932) introduce una rigorosa

impostazione dell’aritmetica basata su tre concetti primitivi e su sei postulati (la formulazione originale della teoria di Peano assume quali concetti primitivi l’unità, il *numero* ed il *successivo*; all’unità lo stesso Peano sostituirà, in un secondo momento, il concetto primitivo *zero*). Illustreremo una costruzione di \mathbb{N} sulla base della teoria degli insiemi.

Gli assiomi di Peano: un approccio ordinale Nella teoria di Peano, così come essa fu definitivamente enunciata in *Aritmetica* (seconda parte del secondo volume di *Formulaire de mathématiques*, 1898), i tre concetti primitivi sono:

1. lo *zero*;
2. il *numero*;
3. il *successivo*.

Gli assiomi sono:

- 1. Assioma zero.** I numeri formano una classe.
- 2. Assioma I.** Lo zero è un numero.
- 3. Assioma II.** Se a è un numero, il suo successivo $a+$ è un numero.
- 4. Assioma III.** Se S è una classe contenente lo zero e, per ogni a , se a appartiene a S , il successivo $a+$ appartiene a S , allora ogni numero naturale è in S (Peano chiama tale proposizione *principio di induzione*).
- 5. Assioma IV.** Se a e b sono due numeri e se i loro successivi $a+$ e $b+$ sono uguali, allora a e b sono uguali.
- 6. Assioma V.** Se a è un numero, il suo successivo $a+$ non è zero.

Nota 5.1. *Sulla necessità dell'Assioma zero, notiamo che esso spiega che alla classe dei numeri naturali possiamo applicare il “calcolo delle classi” che*

Peano stesso aveva sviluppato precedentemente nel proprio libro. Nel nostro linguaggio esso vuol dire che possiamo effettivamente raccogliere i numeri naturali in modo da formare un insieme, cosa che, come abbiamo detto, non è sempre lecita.

La relazione “successivo”, o successore, introdotta da Peano è dunque un’applicazione $s: \mathbb{N} \rightarrow \mathbb{N}$ tale che $s(a) = a+$; si può facilmente dimostrare che, così definita, essa è una funzione iniettiva ma non suriettiva.

Applicando opportunamente gli assiomi, ed approntando le necessarie dimostrazioni, Peano giunse ad introdurre le operazioni aritmetiche con i numeri naturali, l’ordinamento, nonché a descrivere ed a dimostrare le loro proprietà formali (v. più avanti).

5.2 L'induzione

L'Assioma III del sistema di Peano ha un ruolo del tutto particolare in matematica (ed in informatica), perchè esso permette di ragionare su successioni infinite, o quantomeno illimitate, con un metodo finitario. Facciamo qualche esempio.

Supponiamo di voler dimostrare che una certa proprietà P è valida per tutti i numeri naturali. È chiaro che non potremo verificarla in tutti i casi (non termineremmo mai); dovremmo perciò desistere? Non è detto. Con un po' di pratica ci accorgeremmo che le proprietà non sono tutte dello stesso tipo; in particolare ce ne sono di quelle la cui dimostrazione è sostanzialmente la stessa al variare del numero per il quale le verifichiamo e di quelle la cui dimostrazione cambia drasticamente al cambiare del numero. La scomponibilità

in fattori primi è del secondo tipo: come si potrebbe passare ad esempio dalla dimostrazione nel caso 60 a quella nel caso 61? (In realtà questo è possibile ricorrendo ad uno stratagemma che vedremo nella Sezione 5.2.2 e, nel caso specifico, nell’Esercizio 5.10). Un esempio del primo tipo, invece, è la celebre formula dell’aritmetica che fornisce la *somma di tutti i naturali da^o ad n*, con n naturale fissato:

$$S_n = \frac{n \cdot (n + 1)}{2}$$

Nell’espressione precedente sono riassunte infinite somme di naturali, una per ciascun indice naturale n . Verifichiamo la formula proposta in alcuni casi:

$$\begin{aligned} S_0 &= 0 = \frac{0 \cdot 1}{2} \\ S_1 &= 0 + 1 = \frac{1 \cdot 2}{2} \\ S_2 &= 0 + 1 + 2 = \frac{2 \cdot 3}{2} \\ S_3 &= 0 + 1 + 2 + 3 = \frac{3 \cdot 4}{2} \end{aligned}$$

Le considerazioni precedenti provano che la formula proposta è valida fino all'indice $n = 3$; ma appare evidentemente impossibile verificare direttamente *tutte* le (*infinite!*) somme di naturali ottenibili.

Una bozza di dimostrazione di tale risultato può essere impostata nel seguente modo. Sia n un numero naturale. Elenchiamo ordinatamente in una prima tabella di una riga tutti i naturali da 1 a n (possiamo escludere lo 0, la cui addizione non modifica la somma) e in una seconda riga gli stessi naturali considerati in ordine inverso:

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & \dots & n \\ n & n-1 & n-2 & \dots & 2 & 1 \end{array}$$

Se sommiamo in colonna le due righe scritte, otteniamo $n + 1$ per ogni colonna, ovvero n volte il numero $n + 1$. La somma che così si ottiene, e cioè $n \cdot n + 1$, è il doppio della quantità S_n cercata, essendo

stata ottenuta addizionando due volte ciascun naturale compreso tra 1 ed n .¹

Questa dimostrazione elementare fa, però, ricorso ad un passaggio poco rigoroso: l'uso di “...”. Cosa significa? In altre (pseudo)dimostrazioni siamo abituati a dire “e così via”: vuol dire che abbiamo notato una regolarità nella dimostrazione, per la quale si passa da un caso a quello successivo aggiungendo una unità. Ovviamente non possiamo accettare, se non per aiutare l'intuizione, una dimostrazione che contenga l'uso di “...” o di espressioni “e così via”. Chi ci garantisce che le cose andranno sempre così? L'Assioma III di Peano ci permette di risolvere rigorosamente questo problema sia nel caso di dimostrazioni che di definizioni per un'infinità numerabile di casi, qualora esse siano di tipo uniforme rispetto all'indice. Infatti esso ha

portato a formulare un principio generale che va sotto il nome di *Principio di induzione*.

In sostanza esso afferma che se una proprietà vale per il primo numero e, valendo per un certo numero, allora vale anche per il successivo, allora vale per tutti i numeri. Usiamo un tale principio anche nella logica di tutti i giorni quando ci riferiamo all'infinito come in frasi del tipo: “*Il mio bambino sa contare*” che vuol dire “*Tutti i numeri naturali sono conosciuti dal mio bambino*”.

Naturalmente ciò non vuol dire che il mio bambino abbia effettivamente nominato tutti i numeri naturali, e nemmeno li abbia pensati; ma egli è in grado di cominciare la serie e, se qualcuno gli nomina un numero, di dire il successivo. Conosce, cioè, lo schema che permette di nominare qualunque numero a partire dal

precedente. La permanenza dello schema al variare del numero considerato, permette di ridurre gli infiniti passaggi necessarii a due soltanto.

Cerchiamo ora di formalizzare il principio di induzione. Indichiamo dunque con il simbolo $P(n)$ la proposizione da provare per tutti i numeri naturali, sottolineando in tale modo che si tratta di un'affermazione dipendente dall'indice $n \in \mathbb{N}$. Presenteremo ora la dimostrazione per induzione di $P(n)$ in due fasi distinte, *entrambe indispensabili*:

1. prima fase, solitamente chiamata *caso base*: si verifica direttamente la verità di $P(0)$; nel caso in cui la proposizione da dimostrare valga per $n \geq n_0 > 0$, si verifica che essa valga per il minimo degli indici, cioè n_0 .
2. seconda fase, solitamente chiamata

passo induttivo: si ammette la verità di $P(n)$ e, sulla base di ciò, si dimostra che la proposizione P è vera anche per l'indice $n + 1$; ovvero, si prova che la validità della proposizione per un indice (qualsiasi) comporta la validità per l'indice successivo.

Se è possibile completare la verifica di *entrambi* i punti sopra illustrati, la proposizione $P(n)$ può dirsi dimostrata (per *tutti* gli indici $n \in \mathbb{N}$): la prima fase, infatti, ci consente di affermare che la proposizione $P(n)$ è vera per $n = 0$; sulla base di ciò, la seconda fase ci assicura che $P(n)$ è vera anche per $n = 1$ (ovvero per il successivo di 0). Appurato ciò, possiamo affermare che $P(n)$ è vera anche per $n = 2$ (per il successivo di 1) e così di seguito per tutti gli indici naturali.

Il fatto che il principio di induzione sia conseguenza dell'Assioma III è facilmente

provato dal seguente ragionamento. Se S è l'insieme degli indici k per i quali $P(k)$ è vero (risp. definito), allora, una volta provati i due passi dell'induzione (caso base e passo induttivo), avremo dimostrato che $0 \in S$ e, se $n \in S$, allora anche $n+ \in S$; questo, per il suddetto assioma, equivale a dire che S coincide con \mathbb{N} e, cioè, $P(n)$ vale (risp. è definito) per ogni n numero naturale.

Se qualcuno avesse ancora dubbi riguardo alla ragionevolezza del principio di induzione e della sua capacità di raggiungere tutti i casi possibili, può fare il seguente ragionamento, che, si badi bene, non è una dimostrazione, ma una riduzione dello stesso principio al secondo ed al terzo degli assiomi di Campano da Novara, certamente più intuitivi.
Supponiamo, che nonostante si sia provata una certa proprietà P sia per il

primo numero, sia che, una volta supposta per n , sia stata provata per il suo successore, ci sia ancora almeno un numero m per la quale essa non valga. Allora m non può essere 0, perché sappiamo che $P(0)$ è vera per la prima condizione; m sarà quindi il successore di un altro numero che viene prima nella serie. Per questo numero non potrà valere la proprietà P , perché allora varrebbe anche per m , per la seconda condizione. In questo modo, in numero finito di passi potremo regredire fino allo 0 con una successione di elementi che non godono della proprietà P ; ma questo è assurdo, perché P è vera per 0.

5.2.1 Dimostrazioni e definizioni per induzione

Illustriamo il procedimento dimostrativo attraverso un esempio nel quale

renderemo rigorosa la dimostrazione proposta per la formula del precedente paragrafo.

Esempio 5.1. *Sia n un numero naturale. Dimostriamo che la somma S_n di tutti i numeri naturali fino ad n è data dalla formula (già proposta e dimostrata nel precedente esempio):*

$$S_n = \frac{n \cdot (n + 1)}{2}$$

Procediamo per induzione sull'indice n .

1. *Caso base: mostriamo che la formula è verificata per il naturale $n = 0$. Come già visto, $S_0 = 0 = \frac{0 \cdot 1}{2}$.*
2. *Passo induttivo: ammettiamo ora che la formula in questione sia verificata per l'indice n ; ammettiamo cioè che sia vero che*

$$S_n = \frac{n \cdot (n + 1)}{2}$$

Dovremo, sulla base di ciò, provare la validità della formula anche per l'indice $n + 1$. Visto che $S_{n+1} = S_n + (n + 1)$, abbiamo che $S_{n+1} = \frac{n \cdot (n+1)}{2} + n + 1 = \frac{n \cdot (n+1) + 2(n+1)}{2} = \frac{(n+1) \cdot (n+2)}{2}$.

Pertanto, la validità della formula per l'indice n comporta la validità della formula per l'indice $n + 1$. Ciò, unito alla provata validità della formula per $n = 0$, completa la dimostrazione per induzione.

Esempio 5.2. Si può verificare che se l'insieme I è costituito da n elementi (cioè di cardinalità n), l'insieme delle parti di I è costituito da 2^n elementi (ovvero, la cardinalità di $\wp(I)$ è 2^n).

Per esempio, consideriamo l'insieme (avente cardinalità 3) $I = \{5, w, z\}$. I suoi sottoinsiemi propri sono $\{5\}$, $\{w\}$, $\{z\}$, $\{5, w\}$, $\{5, z\}$, $\{w, z\}$; i suoi sottoinsiemi

impropri sono \emptyset e $\{5, w, z\}$. Pertanto, l'insieme delle parti (proprie e improprie) dell'insieme I ha cardinalità $2^3 = 8$.

Tuttavia, questa non è una dimostrazione; per farla su ogni n non si può fissare n ma bisogna procedere per induzione su n .

1. *Caso base: se $n = 0$, allora $I = \emptyset$ e dunque $\#\wp(I) = \#\wp(\emptyset) = 1 = 2^0$. La tesi è quindi valida per questo primo caso.*
2. *Passo induttivo: Ammettiamo la validità della tesi da dimostrare qualora I sia costituito da n elementi, cioè ammettiamo che sia $\#\wp(I) = 2^n$. Sia ora $a \notin I$; consideriamo quindi l'insieme $I \cup \{a\}$ la cui cardinalità è $n + 1$. Qual'è la cardinalità di $\wp(I \cup \{a\})$? Elenchiamo tutti i sottoinsiemi di $\wp(I \cup \{a\})$ in una tabella nel modo seguente: nella prima riga scriviamo tutti i*

sottoinsiemi di $I \cup \{a\}$ che non contengono a (in pratica, tutti e soli i sottoinsiemi di I): $\emptyset = I_1, I_2, I_3, \dots, I_{2n} = I$. In una seconda riga, scriviamo i sottoinsiemi di $I \cup \{a\}$ che contengono a , con un criterio assai semplice: uniamo a ciascun sottoinsieme della prima riga l'insieme $\{a\}$. Otteniamo:

$$\begin{array}{ccccccc} \emptyset & I_2 & I_3 & \dots & I_{2^n-1} & I \\ \{a\} & I_2 \cup \{a\} & I_3 \cup \{a\} & \dots & I_{2^n-1} \cup \{a\} & I \cup \{a\} \end{array}$$

Quanti sono, dunque, i sottoinsiemi di $I \cup \{a\}$? Nella prima colonna ne abbiamo elencati 2^n , in quanto abbiamo ammesso che sia $\wp(I) = 2n$; nella seconda colonna ne troviamo altrettanti. I sottoinsiemi di $I \cup \{a\}$ sono $2 \cdot 2n = 2^{n+1}$ e ciò completa la dimostrazione.

Il lettore faccia attenzione all'esempio seguente.

Esempio 5.3. Vogliamo dimostrare, per induzione, che per ogni numero naturale n sia $(n + 1)^2 - (n - 1)^2 = 4n$. Ammettiamo che la formula sia valida per n , ovvero che sia $(n + 1)^2 - (n - 1)^2 = 4n$, e dimostriamola per $n + 1$. Risulta $(n + 2)^2 - (n)^2 = ((n + 1) + 1)^2 - ((n - 1) + 1)^2 = (n + 1)^2 + 1 + 2(n + 1) - (n - 1)^2 - 1 - 2(n - 1) = (n + 1)^2 + 2n + 2 - (n - 1)^2 - 2n + 2 = (n + 1)^2 - (n - 1)^2 + 4 = 4n + 4 = 4(n + 1)$, che è la formula che si doveva dimostrare.

La precedente dimostrazione è incompleta: manca infatti l'indispensabile verifica della formula per $n = 0$:

$$(0 + 1)^2 - (0 - 1)^2 = 1 - 1 = 0 = 4 \cdot 0$$

La tesi è dunque verificata in questo primo caso. A quest'ultima verifica può essere fatto seguire quanto sopra stabilito: ciò induzione su n . completa la

dimostrazione per

Nota 5.2. *La formula proposta in quest'ultimo esempio avrebbe potuto essere dimostrata anche direttamente, senza ricorrere alla dimostrazione per induzione (sarebbe stato infatti sufficiente sviluppare i quadrati al primo membro).*

Il principio di induzione può essere utilizzato non soltanto per dimostrare proprietà indicizzate sui numeri naturali, ma anche per dare definizioni, qualora queste risultino variare “con regolarità” sui numeri naturali. Anche in questo caso la definizione dovrà essere esibita per un caso base e poi, supponendola data per un caso generico, dovrà essere esibita per il successivo. In questo modo vengono definite le usuali operazioni aritmetiche.

Definizione 5.1. *Fissato un naturale n , la somma con n è così definita:*

$$\begin{aligned}n + 0 &= n \\n + m_+ &= (n + m)_+\end{aligned}$$

Con questa definizione, si può facilmente dimostrare che $n_+ = n_+ \circ_+$, cioè $n_+ = n_+ 1$, che avevamo già usato in maniera informale negli esempi precedenti.

Analogamente, per la moltiplicazione.

Definizione 5.2. *Fissato un naturale n , la moltiplicazione per n è così definita:*

$$\begin{aligned}n \cdot 0 &= 0 \\n \cdot (m_+) &= (n \cdot m) + n\end{aligned}$$

Usando il principio di induzione si possono anche provare per queste operazioni le ben note proprietà commutativa, associativa e di cancellabilità. Il procedimento non è difficile, ma un po' noioso. Vale la pena di provarne effettivamente almeno una come

esercizio: ad esempio, la cancellabilità per la somma, cioè che $n + m = n + m'$ implica $n = n'$, può essere provata per induzione su m .

Una volta definita la somma di numeri naturali, si può introdurre l'ordinamento.

Definizione 5.3. *Dati due numeri naturali, n ed m , si dirà $n \leq m$ se esiste un altro naturale h tale che $n + h = m$; h sarà detta la differenza tra n ed m ed indicata con $m - n$.*

La verifica del fatto che si tratti di un ordinamento viene lasciata come esercizio. Si tratta in realtà di un *buon ordinamento*, perchè, oltre ad essere totale, in esso ogni sottoinsieme non vuoto ha un primo elemento.

Proposizione 5.1. ‘ \leq ’ è un buon ordinamento su \mathbb{N} .

Dimostrazione. Dimostriamo dapprima che i naturali sono totalmente ordinati rispetto a \leq , cioè sono una *catena*. A tale scopo, facciamo vedere che, dati due numeri naturali n ed m , accade sempre che $n \leq m$ o $m \leq n$, cioè sono confrontabili. Sia C l'insieme dei naturali confrontabili con ogni altro. Sicuramente $0 \in C$, perchè $0 + n = n$, per ogni $n \in \mathbb{N}$. Se $m \in C$, allora anche $m + 1 \in C$ e quindi C coincide con \mathbb{N} : infatti, preso comunque n , o $n \leq m$ (cioè $n + h = m$) o $m \leq n$ (cioè $m + k = n$); nel primo caso, $n \leq m + 1$ perchè $n + h + 1 = m + 1$; nel secondo o $k = 0$, $m = n$ e $n \leq m + 1$, oppure $k = r + 1$ e $m + 1 \leq n$. Per esercizio si metta in risalto l'uso dei diversi assiomi di Peano in questa dimostrazione.

Sia dato ora un sottoinsieme non vuoto I di \mathbb{N} . Se $0 \in I$, allora 0 è il minimo di I , perchè ogni numero è maggiore di 0 ;

altrimenti $o \in J$, dove J è l'insieme dei numeri strettamente minori di tutti gli elementi di I . Questo insieme è finito, perchè, non appena un $m \in I$, tutti i numeri da m in poi non possono stare in J , dal momento che \mathbb{N} è totalmente ordinato. I numeri da o ad m sono un insieme finito, perciò minimo di I . J è anch'esso finito ed avrà come massimo n , mentre $n + 1$ è il

□

Questa proprietà dei numeri naturali si chiama *Principio del buon ordinamento* ed è equivalente al principio di induzione (cfr. il terzo assioma di Campano).

5.2.2 Varianti del Principio di Induzione: Induzione strutturale e Induzione completa

Un'apparente generalizzazione del principio di induzione è il seguente *principio di induzione completa*: “Se vale $P(0)$ e, se per ogni $n < m$, la validità di $P(n)$ implica la validità di $P(m)$, allora vale $P(k)$ per ogni k ”. In questo caso l’ipotesi è più debole perché il predecessore di n è soltanto uno dei numeri più piccoli di n ; perciò è evidente che questo principio implica il precedente (la dimostrazione formale si potrà fare come esercizio una volta studiato il calcolo dei predicati). In realtà sui numeri naturali i due principi sono equivalenti.

Esempio 5.4. Si considerino i numeri di Fibonacci $F_0, F_1, F_2, \dots, F_n, \dots$, definiti (*induttivamente*) come segue:

$$F_0 = 0$$

$$F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2} \text{ per ogni } n \geq 2$$

L' n -esimo numero di Fibonacci gode della seguente proprietà: $F^n \leq 2^{n-1}$.

Dimostriamo questa proprietà per induzione (completa) su n . Questo è un tipico caso in cui si hanno due casi base (e non uno soltanto, come abbiamo finora visto), dato che la definizione ricorsiva assume due casi iniziali: $F_0 = 0 \leq 2^{-1} = 1$ e $F_1 = 1 \leq 2^0 = 1$. Per il caso induttivo, supponiamo di aver dimostrato la tesi fino al caso $n - 1$ e dimostriamola per il caso n . Abbiamo che $F_n = F^{n-1} + F^{n-2} \leq 2^{n-2} + 2^{n-3} = 2 \cdot 2^{n-3} + 2^{n-3} = 3 \cdot 2^{n-3} < 4 \cdot 2^{n-3} = 22 \cdot 2^{n-3} = 2^{n-1}$. Si noti che, per legittimare il secondo passaggio, dobbiamo aver dimostrato non soltanto che $F_n^1 \leq 2^{n-2}$, ma anche che $F_n^2 \leq 2^{n-3}$.

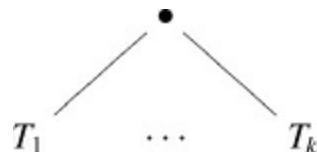
Il principio di induzione può essere applicato, oltre che direttamente all’insieme dei numeri naturali, anche ad un qualunque insieme I equipotente ad esso, cioè numerabile (ossia, “indicizzato” sui numeri naturali), dal momento che esso opera soltanto su indici. Deve però essere esplicitamente fornita la corrispondenza biunivoca tra I ed \mathbb{N} .

Questo ci permetterà, ad esempio, di poter partire da un numero che non sia lo 0. Infatti, per quanto riguarda la prova del caso base, non è necessario che questo venga identificato con lo 0. Immaginiamo una proprietà che valga per i naturali a partire da 25. Il caso base dovrà essere per forza 25, ma ciò non costituisce problema perché potremo etichettare 25 con 0, 26 con 1, ecc.; ossia porre una corrispondenza biunivoca $f: \mathbb{N} \setminus \{0, \dots, 25\} \rightarrow \mathbb{N}$, definendo $f(n) = n - 25$. La prova

per induzione, lavorando sugli indici, avrà
o come caso base.

Abbiamo un altro caso di induzione che
utilizzeremo spesso in logica, che
introdurremo con l'esempio seguente.

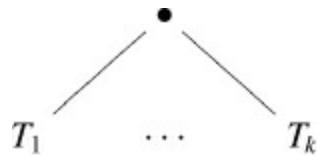
Esempio 5.5. *Diamo una definizione
ricorsiva (ossia, induttiva) di albero finito
non vuoto: un singolo nodo, indicato con
‘•’, è un albero; se T_1, \dots, T_k sono alberi,
allora*



*è un albero (diremo che T_1, \dots, T_k sono figli
del nodo ‘•’, chiamato radice dell’albero).
Definiamo inoltre l’altezza di un albero
(che, talvolta chiameremo altezza del
nodo radice), scritta $h(T)$, come:*

$$h(\bullet) = 0 \quad e \quad h\left(\begin{array}{c} \bullet \\ T_1 & \dots & T_k \end{array} \right) = 1 + \max_{i=1,\dots,k} \{h(T_i)\}$$

Dimostriamo ora che un albero T di altezza n ha al più $(d + 1)^n$ nodi, dove d è il massimo numero di figli di un nodo in T . La dimostrazione è per induzione su $h(T)$. Il caso base è per $h(T) = 0$, cioè T è l'albero formato da un solo nodo (senza figli): allora, il numero di nodi di T è $1 = (0 + 1)^0$, essendo $d = 0$ en = 0. Assumiamo la tesi vera per ogni albero di altezza minore di n e dimostriamola per un generico albero T di altezza n (stiamo quindi usando l'induzione completa!). Abbiamo che T è della forma:



dove $h(T_i) < h(T) = n$, per ogni $i \in \{1, \dots, k\}$; quindi, per induzione, ogni T_i ha al più

$(d_i + 1)n_i$ nodi, dove d_i è il massimo numero di figli di un nodo in T_i e $n_i = h(T_i)$. Sia $d = \max\{k, \max_i \{d_i\}\}$; chiaramente, $(d_i + 1)^{n_i} \leq (d + 1)^{n_i} \leq (d + 1)^{n-1}$, per ogni $i \in \{1, \dots, k\}$. Quindi, il numero di nodi in T è al più

$$\sum_{i=1..k} (d + 1)^{n-1} = k \cdot (d + 1)^{n-1} \leq d \cdot (d + 1)^{n-1} < (d + 1)^n.$$

Questo modo di procedere si dice per *induzione strutturale* ed opera sostanzialmente associando ad oggetti (non necessariamente numeri) che sono decomponibili in oggetti più semplici, un numero naturale che ne rappresenta la complessità, in modo che le componenti abbiano associato un numero più piccolo. La dimostrazione avverrà poi per induzione su questi indici di complessità. Talvolta, quando la diminuzione della complessità sarà evidente, si potrà addirittura omettere di specificare

l'indice.

5.3 Divisibilità e numeri primi

5.3.1 L'algoritmo euclideo e il massimo comun divisore

Introduciamo ora un'importante proprietà dei numeri naturali (che, di fatto, vale anche per i numeri interi). Dati due numeri naturali, n ed m con $m \leq n$ è sempre possibile trovare q ed r con $r < m$ (detti rispettivamente *quoziente* e *resto*) tali che $n = mq + r$. Tali numeri si trovano sottraendo iterativamente m da n , finché ciò che resta è minore di m . Questo procedimento terminerà sicuramente ed il numero delle volte che verrà applicato darà q .

Su questa base si può costruire l'algoritmo

che dia il *massimo comun divisore*, cioè il più grande intero che divida sia n che m . Dati n ed m , si procede calcolando iterativamente il quoziente ed il resto in questa successione:

$$\begin{aligned} n &= mq_0 + r_0 \\ m &= r_0q_1 + r_1 \\ &\dots \\ r_{m-1} &= r_m q_{m+1} \end{aligned}$$

Come spesso capita, il fatto che l’algoritmo termini è basato sul principio di induzione (o del buon ordinamento): infatti, prima o poi, dopo un numero finito di passi, il resto si annullerà, poichè si decrementa ad ogni passo. Lasciamo come facile esercizio la verifica che r_m è il massimo comun divisore tra n ed m . La prova, riportata da Euclide, è forse la prima nella quale si riscontra una tecnica di dimostrazione per induzione, molto prima della formulazione dello stesso

principio.

La possibilità, dati n ed m , di determinare univocamente il resto della divisione, permette di introdurre una relazione di equivalenza tra i numeri stessi.

Definizione 5.4.. *Due numeri n ed n' , si dicono congruenti modulo m , e si scrive $n \equiv n' \pmod{m}$, se, divisi per m danno lo stesso resto.*

Per esercizio, si può verificare che la congruenza modulo un certo numero m è una relazione di equivalenza. In realtà, essa è molto di più, perché le classi che si vengono a formare “ereditano” le operazioni esistenti tra i numeri che le rappresentano; tali equivalenze vengono chiamate *congruenze*. Formalmente, possiamo definire le operazioni tra classi nel seguente modo:

$$[k]_{\equiv_m} + [h]_{\equiv_m} = [k + h]_{\equiv_m} \quad [k]_{\equiv_m} \cdot [h]_{\equiv_m} = [kh]_{\equiv_m}$$

Ciò è lecito (la definizione è ben posta) perchè tale definizione è indipendente dal rappresentante che abbiamo scelto per la classe, ossia si può verificare che, se $k' \equiv k \pmod{m}$ e $h \equiv h' \pmod{m}$, allora

$$[k' + h']_{\equiv_m} = [k + h]_{\equiv_m} \text{ e } [k'h']_{\equiv_m} = [kh]_{\equiv_m}.$$

Naturalmente esistono molte situazioni matematiche esprimibili mediante l'uso di delle congruenze anche al di fuori dei numeri naturali, come vedremo nella seconda parte di questo testo. Tuttavia, le congruenze fanno parte della nostra vita di tutti i giorni. Si pensi al quadrante di un normale orologio a lancette, che scandisce il tempo ciclicamente dividendolo in 12 ore (ignoriamo, per semplicità le lancette dei minuti e dei secondi): ragionare sulle ore è un esempio di operazioni modulo 12.

5.3.2 La rappresentazione dei

numeri naturali

Sulla possibilità di eseguire la divisione con resto è basata la rappresentazione dei naturali in una certa *base*. La moderna rappresentazione dei numeri naturali, ottenibile mediante le dieci cifre 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9, viene detta *posizionale in base dieci*. Ciò significa che il valore di ogni singola cifra che compone la rappresentazione di un numero n dipende dalla posizione di tale cifra in quella rappresentazione; il valore (totale) n del naturale rappresentato da una sequenza ordinata di cifre: $a_m, a_{m-1}, \dots, a_2, a_1, a_0$ è calcolabile mediante l'espressione

$$n = a_m \cdot 10^m + a_{m-1} \cdot 10^{m-1} + \dots + a_2 \cdot 10^2 + a_1 \cdot 10^1 + a_0 \cdot 10^0.$$

Si noti che ciò è legittimato dall'operazione di divisione con resto introdotta precedentemente.

Esempio 5.6. Il valore del numero 35206 (scritto in notazione decimale, cioè in base dieci) è dato dalla somma $3 \cdot 10^4 + 5 \cdot 10^3 + 2 \cdot 10^2 + 0 \cdot 10^1 + 6 \cdot 10^0 = 30000 + 5000 + 200 + 0 + 6$.

Nota 5.3. La storia della matematica registra, lungo il corso dei secoli, il susseguirsi di metodi diversi per la rappresentazione dei naturali. Il sistema di scrittura dei numeri nelle matematiche del mondo antico (con l'eccezione dell'aritmetica babilonese) non si avvale della notazione posizionale: il valore di un numero risulta semplicemente dalla somma dei valori associati ai simboli che, indicati uno di seguito all'altro, vengono a costituire il numero stesso; tali valori sono fissi, cioè non dipendono (a parte rare eccezioni) dalla posizione del simbolo nella scrittura del numero. Una rappresentazione di questo genere per i

naturali, detta additiva, è caratteristica dell’aritmetica romana. In realtà, tale rappresentazione è ulteriormente complicata dal fatto che non è esclusivamente additiva, ma anche sottrattiva: si pensi, ad esempio, alla rappresentazione del numero⁹, cioè IX.

Esempio 5.7. Ricordando che il valore dei simboli romani M , C , L , X , V e I è rispettivamente 1000, 100, 50, 10, 5, 1, il numero romano scritto nella forma $MC\,L\,XXVIII$ è dato, additivamente, da $1000 + 100 + 50 + 10 + 10 + 5 + 1 + 1 + 1 = 1178$.

Tuttavia ogni civiltà ha fissato una base nella quale nominare i numeri. La notazione posizionale ha fatto corrispondere perfettamente i nomi alla rappresentazione, ma per far questo ha avuto bisogno di un simbolo speciale per lo “zero”, che per lungo tempo non è stato

considerato un numero.

Prima di chiudere il presente paragrafo, è doveroso riservare un accenno ai sistemi di numerazione posizionale non decimale, ovvero che si avvalgono di basi diverse da dieci. È infatti possibile scegliere come base del sistema di numerazione, un numero b diverso da dieci (importanti sono, ad esempio per le applicazioni al calcolo automatico, le basi *due* e *sedici*). Ciò significa che il valore n del naturale (in base b) rappresentato dalla sequenza ordinata di cifre (minori di b) $a_m, a_{m-1}, \dots, a_2, a_1, a_0$ è calcolabile mediante l'espressione

$$a_m \cdot b^m + a_{m-1} \cdot b^{m-1} + \dots + a_2 \cdot b^2 + a_1 \cdot b^1 + a_0 \cdot b^0.$$

Esempio 5.8. *Il valore del numero 35206 scritto in notazione posizionale in base 7 è dato dalla somma: $3 \cdot 7^4 + 5 \cdot 7^3 + 2 \cdot 7^2 + 0 \cdot 7^1 + 6 \cdot 7^0$. In base dieci, tale*

numero sarebbe rappresentato da 9022.

Sappiamo che i moderni calcolatori lavorono con una base binaria, molto semplice perchè usa due soli simboli, ma che dà luogo a rappresentazioni praticamente illegibili per gli umani (ciò che è semplice non è sempre facile!).

5.3.3 I numeri primi

In questo paragrafo presenteremo le principali nozioni collegate ai numeri primi e forniremo alcune dimostrazioni. Dapprima ci chiederemo che cosa sono i numeri primi: daremo la definizione, illustreremo il crivello di Eratostene e mostreremo l'esistenza e l'unicità della scomposizione in fattori primi di un numero naturale. Poi ci domanderemo quanti sono i numeri primi e come sono distribuiti nella sequenza dei numeri

naturali. Daremo quindi alcuni risultati: una condizione necessaria di primalità (Teorema di Fermat) e una condizione necessaria e sufficiente di primalità (Teorema di Wilson). Presenteremo infine alcuni problemi aperti, come le congetture di Goldbach e dei primi gemelli.

Iniziamo col presentare la nozione di divisibilità.

Definizione 5.5. *Un naturale a si dice divisibile per un naturale b se esiste un naturale c tale che $a = b \cdot c$. Si dice allora che b è un divisore di a e si scrive $b \mid a$.*

Si tratta di una nozione elementare; ciononostante, su di essa si basano tecniche e concetti di primaria importanza.

Definizione 5.6. *Il naturale p si dice primo se è maggiore di 1 ed è divisibile*

soltanto per 1 e per se stesso. Un naturale maggiore di 1 non primo si dice composto.

Un’antica tecnica per individuare i numeri primi minori di un limite prefissato è illustrata nell’esempio seguente.

Esempio 5.9 (Crivello di Eratostene). *Si vogliono trovare i numeri primi minori di $n \geq 2$. Partiamo con la lista $2, 3, 4, \dots, n$. Iniziamo a prendere in considerazione il naturale 2 ed affermiamo che si tratta di un primo; ne segue che tutti i multipli di 2 saranno composti e non dovranno essere presi in considerazione nel seguito (li cancelliamo dalla lista). Prendiamo ora il numero successivo nella lista, e cioè 3; di nuovo, esso è primo e tutti i suoi multipli vanno cancellati dalla lista. Ora, il numero*

successivo nella lista è 5 (poichè 4 è stato cancellato, essendo multiplo di 2); e così via. Il procedimento deve essere ripetuto considerando successivamente tutti i naturali fino ad arrivare a \sqrt{n} (Perchè? Il lettore è invitato a rispondere per iscritto, per esercizio).

Passiamo ora a formulare alcune cruciali proprietà dei numeri naturali: l'esistenza ed unicità della *scomposizione in fattori primi*, e il fatto che ci sia un'infinità di numeri primi.

Proposizione 5.2 (Esistenza della scomposizione in fattori primi).
Ogni numero naturale maggiore di 1 è un prodotto di numeri primi.

Dimostrazione. Sia n un numero naturale; o n è primo, nel qual caso la tesi è provata, oppure n ha divisori compresi tra 1 e n . Sia p_1 il minimo di questi

divisori; p_1 è primo altrimenti, se ammettesse un divisore k compreso tra 1 e p_1 stesso, k sarebbe divisore anche di n , contro la minimalità di p_1 . Quindi $n = p_1 \cdot n_1$, con $1 < n_1 < n$. Ripetiamo il ragionamento su n_1 : o esso è primo o è divisibile per un primo $p_2 < n_1$. Iterando il procedimento, otteniamo una sequenza decrescente di numeri (non primi) n, n_1, n_2, \dots, n_k , che terminerà, essendo $n_k + 1 = p_{k+1}$ un numero primo. Allora $n = p_1 \cdot p_2 \cdot \dots \cdot p_{k+1} \cdot$

□

Proposizione 5.3 (Unicità della scomposizione in fattori primi). *La scomposizione in fattori primi di un numero naturale è unica, a meno di permutazioni di fattori.*

Dimostrazione. (Lindemann) Chiamiamo

numeri anormali i numeri che possono essere fattorizzati in prodotti di primi in più modi (a meno di permutazioni). Sia n il minimo numero anormale. Osserviamo anzitutto che lo stesso primo p non può apparire in due diverse fattorizzazioni di n : se così fosse, n / p sarebbe un numero anormale ma minore di n , contro la minimalità di n . Ora sia $n = p_1 p_2 p_3 \dots = q_1 q_2 q_3 \dots$, dove i p_i e i q_j sono primi e per ogni i e j $p_i \neq q_j$.

Senza perdere di generalità, assumiamo che p_1 e q_1 siano il minimo dei p_i e dei q_j ; risulta $p_1 < n$ e $q_1 < n$. Da ciò, ricordando che $p_1 \neq q_1$, si ha che $p_1 q_1 < n$. Poniamo $m = n - p_1 q_1$; essendo $0 < m < n$, non può essere che m sia un numero anormale. Ora, $p_1 | n$ e dunque $p_1 | m$; similmente, $q_1 | n$ e dunque $q_1 | m$. Ciò significa che p_1 e q_1 appaiono entrambi

nell’(unica) fattorizzazione di m e quindi $p_1q_1 \mid m$.

Da questo segue che $p_1q_1 \mid n$ e quindi che $q_1 \mid n / p_1$. Ma n / p_1 è $p_2 p_3 \dots$ che, essendo minore di n , ammette un’unica fattorizzazione. Dato che q_1 non è uno dei p_i , questo è impossibile. Dunque non possono esistere numeri anormali.

□

Proposizione 5.4 (XX, Libro IX degli Elementi). *I numeri primi sono sempre più di ogni assegnata quantità di primi.*

Dimostrazione. Sia p_1, p_2, \dots, p_r un’assegnata quantità di numeri primi. Poniamo $n = p_1 p_2 \dots p_r + 1$; se n è primo, abbiamo concluso. Sia quindi p un numero primo che divida n ; allora p non può essere alcuno dei p_1, p_2, \dots, p_r altrimenti p dividerebbe $n - p_1, p_2, \dots, p_r =$

1, che è impossibile. Dunque, p è un altro primo.

□

Enunciamo ora alcune classiche condizioni di primalità. Una condizione necessaria affinchè un numero naturale sia primo è espressa dal *Piccolo Teorema di Fermat* (dimostrato da Eulero). Esso dice che, se p è un primo, allora $a_p - a$ deve essere un multiplo di p , qualunque sia a .

Proposizione 5.5. *Se p è un numero primo, allora $a p - a \equiv 0 \pmod{p}$ per ogni $a \in \mathbb{Z}$.*

La precedente condizione è necessaria, ma non sufficiente: dunque tutti i numeri primi certamente soddisfano quanto espresso dal Piccolo Teorema di Fermat, ma non tutti i numeri che soddisfano tale

condizione sono primi. Una condizione necessaria e sufficiente è espressa dal *Teorema di Wilson* (introdotto nel 1770, poi dimostrato da Lagrange e da Euler):

Proposizione 5.6. *p è un numero primo se e solo se $(p - 1)! + 1 \equiv 0 \pmod{p}$.*

Può essere interessante fare qualche verifica. Occupiamoci ad esempio del numero primo 7: si ha che $(7 - 1)! + 1 = 721$ è un multiplo di 7 (infatti, $721 = 7 \cdot 103$); ma già la verifica relativa ai primi successivi (per esempio, 11 o 13) appare difficoltosa per il calcolo del fattoriale (il lettore provi a calcolare $10!$ e $12!$). Nel caso di applicazione a primi più grandi, il calcolo del fattoriale si rivela un ostacolo molto serio.

Il Teorema di Wilson quindi è un risultato elegante e importante, ma praticamente ben poco utile, in quanto

non conosciamo alcun algoritmo in grado di calcolare il fattoriale con “sufficiente” rapidità. Da ciò segue che la velocità di un test di primalità basato sul Teorema di Wilson sarebbe minore di quella di un test basato sul crivello di Eratostene. Molte formule sono basate sul Teorema di Wilson; ad esempio la formula di Willans (1964) fornisce il numero primo n -esimo:

$$p_n = 1 + \sum_{k=1}^{2^n} \left[\left(\frac{n}{\sum_{j=1}^k \cos^2 \pi \frac{(j-1)!+1}{j}} \right)^{\frac{1}{n}} \right]$$

ma le difficoltà applicative precedentemente menzionate restano.

Ricordiamo che nella storia dell’informatica le verifiche di primalità sono stati i primi procedimenti ad essere condotti su elaboratori. E tuttora sono considerati validissimi test per valutare l’efficienza di una macchina.

Esistono anche numerosi problemi

aperti sui numeri primi. Uno dei più celebri è la *congettura di Goldbach*, suggerita (ma non provata) in uno scambio di lettere tra Christian Goldbach (1690-1764) e Leonhard Euler (1707-1783) nel giugno 1742, secondo la quale

tutti i naturali pari maggiori di 2 sono somme di due numeri primi (non necessariamente distinti).

La congettura di Goldbach è facilmente verificabile per alcuni pari: ad esempio, $4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $10 = 3 + 7 = 5 + 5$, etc. Ma ciò vale per *tutti* i numeri pari? Nessuno finora (2008) è stato capace di dimostrarlo nè di trovare un naturale pari maggiore di 2 che *non* sia esprimibile come somma di due numeri primi.

Un'altra celebre congettura a tutt'oggi non provata è detta *dei primi gemelli*.

Alcune coppie di numeri primi sono costituite da p e da $p + 2$ (primi gemelli): ad esempio, 3 e 5, 11 e 13, 17 e 19, etc.

Ebbene, esistono infinite coppie di primi gemelli? Ancora una volta il problema è aperto (2008).

Nota 5.4. *La presenza di questi e di altri problemi aperti riguardanti i primi (anche apparentemente semplici: esistono infiniti primi della forma $n_2 + 1$?) induce a qualche riflessione. C'è una forte asimmetria tra l'addizione e la moltiplicazione. Rispetto alla moltiplicazione (elemento neutro 1), infatti, esistono infiniti elementi “atomici”: i primi. La scomposizione di un naturale in naturali “atomici rispetto alla moltiplicazione” è interessante e semplifica la trattazione del numero dato. Invece rispetto all'addizione (elemento neutro 0) esiste un solo elemento atomico:*

1. La scomposizione di un naturale in un prodotto di naturali “atomici rispetto all’addizione” è banale ($1 + 1 + \dots + 1$) e non semplifica la trattazione del numero dato.

Esercizi svolti

Esercizio 5.1. *Dimostrare in tutti i modi possibili che, per ogni $n \in \mathbb{N}$, vale $(n + 1)^2 - n^2 = 2n + 1$.*

Soluzione: Un primo modo è il modo algebrico: $(n + 1)^2 - n^2 = n^2 + 1 + 2n - n^2 = 2n + 1$. Un secondo modo è usando il principio di induzione (che però in questo caso non è la strategia di dimostrazione migliore). Il caso base è per $n = 0$: $(0 + 1)^2 - 0^2 = 1^2 = 1 = 2 \cdot 0 + 1$. Assumendo che $(n + 1)^2 - n^2 = 2n + 1$, dobbiamo dimostrare che $((n + 1) + 1)^2 - (n + 1)^2 = 2(n + 1) + 1$.

Usando semplici passaggi algebrici e l'ipotesi induttiva (abbiamo sottolineato il posto in cui l'abbiamo usata), si ha che

$$((n+1)+1)^2 - (n+1)^2 = \underline{(n+1)^2} + 1 + 2(n+1) - \underline{n^2} - 2n - 1 = \underline{2n+1} + 1 + 2(n+1) - 2n - 1 = 2(n+1) + 1.$$

Esercizio 5.2. *Dimostrare che $\sum_{i=0}^n 2^i = 2^{n+1} - 1$, per ogni $n \in \mathbb{N}$.*

Soluzione: Il caso base è per $n = 0$ e abbiamo che $\sum_{i=0}^0 2^i = 2^0 = 1 = 2^1 - 1$. Per il passo induttivo, supponiamo la tesi vera per n e cerchiamo per $n + 1$. Abbiamo che

$$\sum_{i=0}^{n+1} 2^i = \underline{\sum_{i=0}^n 2^i} + 2^{n+1} = \underline{2^{n+1} - 1} + 2^{n+1} = 2 \cdot 2^{n+1} - 1 = 2^{n+2} - 1.$$

Esercizio 5.3. *Dimostrare per tutti gli $n \in \mathbb{N}$ che $n < 2^n$.*

Soluzione: Il caso base si ha per $n = 0$, per cui abbiamo $0 < 2^0 = 1$. Assumiamo la tesi vera per n e dimostriamola per $n + 1$. Si ha

che $\underline{n} + 1 < \underline{2^n} + 1 \leq 2_n + 2^n = 2^{n+1}$; infatti, si ha che $2^n \geq 1$ per ogni $n \in \mathbb{N}$ (questo può essere sia dimostrato per induzione su n , che con ragionamenti analitici: la funzione esponenziale di base maggiore di 1 è strettamente crescente e $2^0 = 1$. Anche in questo caso il ragionamento induttivo non è il più semplice).

Esercizio 5.4. *Dimostrare che per un sottoinsieme infinito di \mathbb{N} , cioè a partire da un opportuno caso base, $n_2 < 2^n$.*

Soluzione: Sembra che il sottoinsieme cercato possa essere tutto \mathbb{N} . Infatti, come caso base, o soddisfa la diseguaglianza. Proviamo ora a dimostrare il passo induttivo:

$(n + 1)^2 = \underline{n^2} + 2n + 1 < \underline{2^n} + 2n + 1$; se fosse $2n + 1 \leq 2^n$, potremmo concludere che $2^n + 2^n + 1 \leq 2^n + 2^n = 2^{n+1}$, come richiesto. Il problema

è che non per ogni n vale $2^n + 1 \leq 2^n$: per esempio, per $n = 2$ ciò non vale!

Cerchiamo quindi un altro caso base. Andando per tentativi, proviamo con $n = 5$: in questo caso, abbiamo $5^2 = 25 < 2^5 = 32$. Per quanto riguarda il passo induttivo, notiamo che, per $n \geq 5$, vale $2^n + 1 < n_2$ (questo può essere visto con procedimenti analitici o per induzione); quindi, usando due volte l'ipotesi induttiva, abbiamo che $(n+1)^2 = \underline{n^2} + 2n + 1 < \underline{2^n} + 2n + 1 < 2^n + \underline{n^2} < 2^n + \underline{2^n} = 2^{n+1}$.

Pertanto un possibile sottoinsieme di \mathbb{N} per cui la diseguaglianza vale è $\{x \in \mathbb{N}: x \geq 5\}$.

Esercizio 5.5. Sia $fm: \mathbb{N} \rightarrow \mathbb{N}$ una famiglia di funzioni definite come $fm(0) = m$ e $fm(n+1) = fm(n) + 1$, per ogni $m \in \mathbb{N}$. (a) Qual è il risultato di $f_{52}(39)$? (b) Quale proprietà di $fx(y)$ si può dimostrare

in generale?

Soluzione: Prima di rispondere alla domanda (a), cerchiamo di rispondere alla (b); in questo modo, la risposta alla (a) sarà banale.

Con un po' di attenzione, si può notare che la funzione $f_m(n)$ non fa altro che sommare n a m , incrementando di 1 il numero m per n volte. Andiamo quindi a dimostrare che, per ogni $n \in \mathbb{N}$, si ha $f_m(n) = m + n$. Il caso base è per $n = 0$ e, banalmente, $f_m(0) = m = m + 0$.

Assumiamo vera la tesi fino a n e dimostriamola per $n + 1$: si ha che $f_m(n + 1) = \underline{f_m(n)} + 1 = \underline{m + n} + 1 = m + (n + 1)$.

Quindi, $f_{52}(39) = 91$.

Esercizio 5.6. *Dimostrare che, per tutti i numeri naturali, $n(n_2 + 5)$ è divisibile per 6.*

Soluzione: Il caso base è per $n = 0$ e abbiamo che $0(0^2 + 5) = 0$, che è divisibile per ogni numero. Assumiamo la tesi vera fino a n e dimostriamola per $n + 1$.

Abbiamo che $(n + 1)((n + 1)^2 + 5) = n((n + 1)^2 + 5) + (n + 1)^2 + 5 = n(n^2 + 1 + 2^n + 5) + (n + 1)^2 + 5 = n(n^2 + 5) + n(1 + 2n) + (n + 1)^2 + 5 = n(n^2 + 5) + 3n(n + 1) + 6$; basta dimostrare che $n(n^2 + 5)$ e $3n(n + 1)$ sono divisibili per 6. Infatti, per dimostrare che $x + y$ è divisibile per z , è sufficiente dimostrare che sia x che y sono divisibili per z . Ovviamente, 6 è divisibile per 6. Per induzione, $n(n^2 + 5)$ è divisibile per 6. Per quanto riguarda $3n(n + 1)$ distinguiamo due casi: se n è pari (e quindi è della forma $2m$), allora $3n(n + 1) = 6m(2m + 1)$ ed è divisibile per 6; altrimenti, n è della forma $2m + 1$ e $3n(n + 1) = 3(2m + 1)(2m + 2) = 6(2m + 1)(m + 1)$, che è divisibile per 6.

Esercizio 5.7. Sia \mathcal{L} l'insieme di parole sull'alfabeto $A = \{a, b\}$ definito da:

1. $a \in \mathcal{L}$;
2. se $w \in \mathcal{L}$ allora $wx \in \mathcal{L}$, con $x \in A$.

Si dimostri che \mathcal{L} è il linguaggio delle parole che cominciano con a .

Soluzione: Formalmente, un *alfabeto* A è un insieme di caratteri; una *stringa* w è una sequenza di caratteri dell'alfabeto (N.B.: l'ordine con cui appaiono i caratteri conta e sono ammesse ripetizioni di uno stesso carattere!); un *linguaggio* \mathcal{L} è un insieme di stringhe costruite a partire da un dato alfabeto. L'insieme di tutte le stringhe costruite a partire dall'alfabeto A è denotato A^* .

Esplicitando la definizione ricorsiva di \mathcal{L} , abbiamo che $\mathcal{L} = \{a, aa, ab, aaa, aab, aba, abb, \dots\}$. Per dimostrare la tesi,

dobbiamo dimostrare che $\mathcal{L} = \{w \in A^*: w = aw', \text{ per qualche } w' \in A^*\}$. La prova è per doppia inclusione.

Iniziamo col dimostrare che \mathcal{L} contiene soltanto parole che iniziano con a . Sia w una stringa di \mathcal{L} ; per induzione sulla lunghezza di w (cioè, sul numero di caratteri che formano la stringa), dimostriamo che w inizia con a ; data l'arbitrarietà di w , questo basta a dimostrare la tesi. Il caso base si ha quando w è formata da un solo carattere (infatti, \mathcal{L} non contiene la stringa vuota, scritta ε , cioè l'unica stringa lunga 0): in questo caso, l'unica stringa lunga 1 in \mathcal{L} è la stringa formata dal solo carattere a .

Assumiamo la tesi vera per tutte le stringhe di lunghezza $n (\geq 0)$ e dimostriamo la tesi per un'arbitraria stringa w lunga $n + 1$. In base alla definizione ricorsiva del linguaggio, $w = w'$

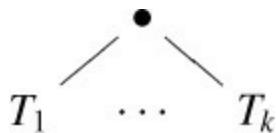
' x , per $w' \in \mathcal{L}$ e $x \in A$. Ovviamente, w è una stringa di \mathcal{L} lunga n , per cui essa inizia con a (per ipotesi induttiva); allora, banalmente, anche w inizia con a (infatti, il primo carattere di w coincide con il primo carattere di w').

Ora dobbiamo dimostrare che ogni stringa che cominci con a è contenuta in \mathcal{L} . Se w inizia con a , allora $w = aw'$, per qualche $w' \in \{a, b\}^*$; la dimostrazione è per induzione sulla lunghezza di w' ed è lasciata per esercizio al lettore.

Esercizio 5.8. *Dimostrare per induzione che in un albero finito e non vuoto T il numero di nodi (che indicheremo con $\#_n(T)$) supera di 1 il numero degli archi (che indicheremo con $\#_a(T)$).*

Soluzione: Dimostriamo la tesi per induzione sull'altezza dell'albero T . Il caso base è per l'(unico) albero di altezza 0,

l’albero ‘•’, che ha un nodo e nessun arco. Assumiamo la tesi vera per ogni albero di altezza al più n e dimostriamola per un generico albero T di altezza $n + 1$. Abbiamo che T è della forma:



dove $h(T_i) < h(T) = n + 1$, per ogni $i \in \{1, \dots, k\}$; quindi, per induzione, $\#_n(T_i) = \#_a(T_i) + 1$. Allora,

$$\begin{aligned} \#_n(T) &= 1 + \sum_{i=1}^k \#_n(T_i) = 1 + \sum_{i=1}^k (\#_a(T_i) + 1) = \\ &= 1 + (k + \sum_{i=1}^k \#_a(T_i)) = 1 + \#_a(T). \end{aligned}$$

Esercizio 5.9. Sia $A \subseteq \mathbb{N}$; se $\mathbb{N} - A$ è finito allora A ed \mathbb{N} sono equipotenti.

Soluzione: Procediamo per induzione sul numero di elementi in $\mathbb{N} - A$. Il caso base si ha per $|\mathbb{N} - A| = 0$; in questo caso, abbiamo che $A = \mathbb{N}$ (infatti, $|\mathbb{N} - A| = 0$ vuol

dire che $\mathbb{N} - A = \emptyset$ e quindi $A = \mathbb{N}$, essendo $A \subseteq \mathbb{N}$), e banalmente A e \mathbb{N} sono equipotenti. Sia la tesi vera fino a n e dimostriamola per

$|\mathbb{N} - A| = n + 1$; questo vuol dire che A contiene almeno un elemento k e sia $A' = A \setminus \{k\}$. Abbiamo che $|\mathbb{N} - A'| = n$ e, per induzione, possiamo trovare una biiezione $f: A \rightarrow \mathbb{N}$. Definiamo la funzione $f': A' \rightarrow \mathbb{N}$ in modo che

$$f'(x) = \begin{cases} n & \text{se } x = n \\ f'(x) & \text{se } f'(x) < n \\ f'(x) + 1 & \text{altrimenti} \end{cases}$$

Non è difficile mostrare che tale funzione è iniettiva e suriettiva poichè f' lo è. Abbiamo quindi trovato una biiezione tra A e \mathbb{N} , dimostrando così che tali insiemi sono equipotenti.

Esercizio 5.10. *Dimostrare per induzione che ogni numero è*

scomponibile in un prodotto di fattori primi.

Soluzione: In questo caso l'induzione non può essere applicata direttamente ai numeri in quanto tali, perché nel passaggio da un numero al successore la situazione e la dimostrazione cambiano drasticamente. Dovremo allora trattare i nostri numeri come oggetti ai quali è assegnato un altro numero che deve variare in modo da rendere parametrica la dimostrazione. Possiamo pensare di associare ad ogni numero un “albero di scomposizione”, cioè etichettare con il numero la radice dell’albero e, se possibile, generare due figli, usando una possibile scomposizione non banale del numero. Ad esempio



Abbiamo due possibilità: o il numero è già primo e ci fermiamo al primo nodo (radice), o si può scomporre in due fattori diversi da 1 e da lui stesso; a questo punto il gioco si ripete per i due fattori e deve terminare prima o poi perchè i fattori ogni volta sono strettamente minori del numero dato e perciò, prima o poi ci fermeremo.

Proviamo ora che qualunque sia l'altezza dell'albero così ottenuto, il numero dato è il prodotto dei numeri primi che sono sulle *foglie* (cioè i nodi senza figli). Se dunque l'albero associato al numero ha altezza 0 vuol dire che il numero dato era già primo e siamo arrivati. Supponiamo che la proprietà sia vera per tutti numeri che hanno un albero di scomposizione di altezza al più h e dimostriamola per un numero n che ha un albero di scomposizione di altezza $h + 1$.

Per un tale numero, esistono due numeri più piccoli r ed s tali che $n = r \cdot s$ ed essi hanno un albero di scomposizione di altezza al più n . Quindi, vale il fatto che essi sono uguali al prodotto dei numeri primi che sono sulle rispettive foglie; ma allora n sarà uguale al prodotto delle due scomposizioni.

Esercizio 5.11. *Dimostrare che, dati tre numeri naturali non nulli tali che la differenza tra il secondo ed il primo e la differenza tra il terzo ed il secondo sia 2, uno di essi è divisibile per 3.*

Soluzione: Siano $n, n + 2, n + 4$ i tre naturali in questione, con $n \neq 0$. La dimostrazione può essere scritta sinteticamente utilizzando le *congruenze* (ricordiamo che $a \equiv b \pmod{n}$ significa che n divide $b - a$):

- se $n \equiv 0 \pmod{3}$, allora la tesi è subito provata;
- se $n \equiv 1 \pmod{3}$, allora $n + 2 \equiv 0 \pmod{3}$;
- se $n \equiv 2 \pmod{3}$, allora $n + 4 \equiv 0 \pmod{3}$.

Altrimenti è necessario esplicitare il ragionamento. Se n è divisibile per 3, la tesi è subito provata. Se n non è divisibile per 3, effettuando tale divisione avremo un resto non nullo r che potrà essere 1 oppure 2. Se $r = 1$, allora $n + 2$ è divisibile per 3: infatti, $n = 3k + 1$ e, pertanto, $n + 2 = 3k + 3 = 3(k + 1)$. Se $r = 2$, allora $n + 4$ è divisibile per 3: infatti, $n = 3k + 2$ e, pertanto, $n + 4 = 3k + 6 = 3(k + 2)$.

Esercizio 5.12. *Dimostrare che, presi due naturali non nulli n ed m , n è divisore di m e m è divisore di n se e solo se $n = m$.*

Soluzione: Chiaramente, se $n = m$ allora sia n è divisore di m che viceversa. Se n è

divisore di m e m è divisore di n avremo che $m = b \cdot n$ e $n = a \cdot m$, per opportuni a e b diversi da zero (essendo, per ipotesi, diversi da zero n ed m). Allora, $n \cdot m = (a \cdot m) \cdot (b \cdot n) = a \cdot b \cdot m \cdot n$; questo implica che $a \cdot b = 1$. Nei naturali, questo è verificato solamente nel caso in cui $a = b = 1$. Questo basta a concludere.

Esercizio 5.13. *Si dimostri che se un numero primo p è divisore di un prodotto $a \cdot b$, allora p deve dividere a o b . Cosa si può dire se p non è primo?*

Soluzione: Siano $p_1 p_2 p_3 \dots$ e $q_1 q_2 q_3 \dots$ le (uniche) scomposizioni in fattori primi di a e di b ; chiaramente, $ab = p_1 p_2 p_3 \dots q_1 q_2 q_3 \dots$. Poichè la scomposizione in fattori primi di ab è unica (a meno di permutazioni dei fattori), deve essere che p è uno dei p_i o uno dei q_j ; pertanto, p

deve dividere a o b .

Se p non è primo, può capitare che non divida né a né b . Si consideri, come controesempio, 12 come prodotto di 3 e 4: 6 divide 12 ma non divide né 3 né 4.

Esercizio 5.14. *Sia p un numero naturale primo dispari. Si dimostri che p può essere espresso come differenza dei quadrati di due naturali in un unico modo.*

Soluzione: Iniziamo col notare che, se p è un primo dispari, allora $\frac{p+1}{2}$ e $\frac{p-1}{2}$ sono numeri naturali tali che

$$\left(\frac{p+1}{2}\right)^2 - \left(\frac{p-1}{2}\right)^2 = \frac{p^2 + 2p + 1 - p^2 + 2p - 1}{4} = p.$$
 Per l'unicità di tale scrittura (a meno di commutatività), osserviamo che, se $p = n_2 - m_2 = (n + m)(n - m)$, allora, essendo 1 e p gli unici divisori di p , deve essere $n - m = 1$ e $n + m = p$, da cui $\frac{p+1}{2}$ e $m \frac{p-1}{2}$.

Esercizio 5.15. Possono esistere due numeri naturali quadrati (diversi da zero) a e b tali che sia $a^2 = 45b^2$? Se sì, fare un esempio; se no, dimostrare l'impossibilità.

Soluzione: Non è possibile. Ammettiamo infatti per assurdo che esistano i naturali n ed m tali che $n^2 = 45m^2$, cioè tali che $n^2 = 3^2 \cdot 5 \cdot m^2$. Consideriamo le scomposizioni in fattori primi: nella scomposizione di n^2 è presente un numero pari di fattori 5 (essendo un quadrato); invece nella scomposizione del numero $3^2 \cdot 5 \cdot m^2$ è presente un numero dispari di fattori 5. Per l'unicità della scomposizione in fattori primi, possiamo concludere.

Esercizio 5.16. Si scelgano k numeri tra $\{1, 2, \dots, 2^n\}$, per un qualche n . Se $k \geq n + 1$, si dimostri che, comunque vengano scelti i k numeri, è impossibile che essi abbiano

un divisore comune.

Soluzione: Comunque si scelgano più di n numeri dagli interi positivi non maggiori di 2^n , c'è almeno una coppia di essi costituita da interi consecutivi: soltanto scegliendone al più n potremmo infatti evitare tale consecutività. Notiamo che due interi consecutivi a ed $a + 1$ non possono avere un divisore $d > 1$ comune: se così non fosse, d dovrebbe dividere anche la differenza tra $a + 1$ ed a , cioè 1, e questo sarebbe assurdo.

Esercizi da svolgere

Esercizio 5.17. *Dimostrare in tutti i modi possibili che, per ogni n naturale, si ha $(n + 2)^2 - n^2 - 4 = 4n$.*

Esercizio 5.18. *Dimostrare che $1 + 3 + 5 + \dots + (2n - 1) = n^2$.*

$$+ \dots + (2n - 1) = n^2.$$

Esercizio 5.19. Dimostrare che $1 + 5 + 9 + \dots + (4n + 1) = (n + 1)(2n + 1)$.

Esercizio 5.20. Si dimostri per induzione che $\sum_{i=1}^n \frac{1}{\sqrt{k}} > \sqrt{n}$; qual è il minimo caso base ammissibile?

Esercizio 5.21. Sia $a \neq 1$; provare per induzione che, per ogni $n \in \mathbb{N} \setminus \{0\}$, si ha che $n \sum_{i=1}^n a^i = a^{\frac{1-a^n}{1-a}}$.

Esercizio 5.22. Provare per induzione che, se a è un reale positivo, allora $(1 + a)n \geq 1 + na$, per ogni $n \in \mathbb{N}$.

Esercizio 5.23. Siano x e y due naturali distinti e non nulli; si dimostri per induzione che, per ogni $n \in \mathbb{N}$, $x - y$ divide $x^n - y^n$.

Esercizio 5.24. Provare per induzione che $\prod_{k=2}^n \left(1 - \frac{1}{k}\right) = \frac{1}{n}$ (N.B.: la notazione \prod ... denota il prodotto, così come la notazione... Σ denota una somma); qual è il minimo caso base ammissibile?

Esercizio 5.25. Dimostrare che, per tutti i numeri naturali non nulli, $5n + 2 \cdot 3^{n-1} + 1$ è divisibile per 8.

Esercizio 5.26. Qual è il più grande sottoinsieme A di \mathbb{N} tale che $n^3 \geq n + 6$, per ogni $n \in A$? Si giustifichi la risposta con un procedimento induttivo.

Esercizio 5.27. Dimostrare che in un ricevimento in cui sono presenti m persone, se ognuno stringe una sola volta la mano di ciascuna altra persona presente, si hanno esattamente $\frac{m(m-1)}{2}$ strette di mano.

Esercizio 5.28. Si consideri il linguaggio \mathcal{L} sull'alfabeto $\{a, b\}$ tale che

- la stringa vuota, ε , appartiene a \mathcal{L} ;
- se $w \in L$ allora anche awa e bwb appartengono a \mathcal{L} .

Si dimostri per induzione matematica che tutte le stringhe in \mathcal{L} sono palindromi (cioè sono uguali se lette da sinistra o da destra).

Esercizio 5.29. Si definisca la funzione $S: \mathbb{N} \rightarrow \mathbb{N}$ tale che $S(0) = 1$ e $S(k + 1) = S(k) + 2^{k+1}$, per ogni $k \in \mathbb{N}$. Si dimostri che, per ogni $n \in \mathbb{N}$, vale che $2^{n+1} - S(n) = 1$.

Esercizio 5.30. Definiamo il grado $d_T(x)$ di un nodo x in un albero T come il numero di archi che toccano x . Sia $\mathbb{N}_{od}(T)$ l'insieme dei nodi di T ; si dimostri che, se $|Nodi(T)| = n$, allora $\sum_{x \in Nodi(T)} d_T(x) = 2n$

– 2.

Esercizio 5.31. *Dimostrare che la somma di cinque numeri naturali consecutivi è sempre divisibile per 5.*

Esercizio 5.32. *Siano a e b due naturali multipli del naturale $n \neq 0$; dimostrare che ogni naturale della forma $\alpha a + \beta b$, con $\alpha, \beta \in \mathbb{N}$, è anch'esso multiplo di n .*

Esercizio 5.33. *Si consideri la seguente rappresentazione tabellare di \mathbb{N} :*

| | | | | | |
|-----------|----------|----------|-------|-----------|-------|
| 1 | 2 | 3 | 4 | 5 | 6 |
| 7 | 8 | 9 | 10 | 11 | 12 |
| 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 |
| 25 | 26 | 27 | 28 | 29 | 30 |
| | | | | | |

dove in grassetto sono evidenziati i numeri primi. Si discuta il perchè, a parte per la prima riga, i primi sono contenuti nella prima e nella quinta colonna.

Esercizio 5.34. *Nel piano cartesiano, chiamiamo punti primi i punti (m, n) aventi entrambe le coordinate appartenenti all'insieme dei numeri (naturali) primi. Si dimostri che nessuna retta passante per l'origine degli assi, ad eccezione della bisettrice del primo quadrante, può passare per più di un punto primo.*

Esercizio 5.35. *Si dimostri che non possono esistere due numeri naturali quadrati (diversi da zero) che siano uno il doppio dell'altro.*

Esercizio 5.36. *Dati i naturali a e b , dimostrare che il numero naturale $a^3 b - ab^3$ è divisibile per tre.*

Esercizio 5.37. *L'insieme $\{m \in \mathbb{N} : m = n^2 + n + 41, \text{ per } n \in \mathbb{N}\}$ contiene soltanto numeri primi? Si giustifichi la risposta.*

¹ Il procedimento illustrato è ricordato in relazione ad un aneddoto riguardante Carl Friedrich Gauss (1777-1855) che, fanciullo, durante un'esercitazione scolastica calcolò velocemente la somma dei naturali da 1 a 100 giungendo a $S_{100} = \frac{100 \cdot 101}{2} = 5050$.

Algebre di Boole

6.1 Proprietà di operazioni aritmetiche e insiemistiche

Un confronto tra le operazioni aritmetiche di addizione e di moltiplicazione da un lato e le operazioni insiemistiche di unione e di intersezione dall'altro ci consentirà di evidenziare analogie (e differenze) che potranno essere utilmente riprese in studi ulteriori. Abbiamo già visto che le operazioni aritmetiche di

addizione e di moltiplicazione su \mathbb{N} godono delle proprietà *associativa* e *commutativa*; cioè per ogni scelta dei numeri naturali n, m e k risulta:

- $(n + m) + k = n + (m + k)$ (proprietà associativa dell'addizione);
- $n + m = m + n$ (proprietà commutativa dell'addizione);
- $(nm)k = n(mk)$ (proprietà associativa della moltiplicazione);
- $nm = mn$ (proprietà commutativa della moltiplicazione).

Non sarà inutile ribadire che anche le operazioni insiemistiche di unione e di intersezione fra sottoinsiemi di un insieme X godono di analoghe proprietà, come abbiamo precedentemente visto:

- $(A \cup B) \cup C = A \cup (B \cup C)$ (proprietà associativa dell'unione);

- $A \cup B = B \cup A$ (proprietà commutativa dell'unione);
- $(A \cap B) \cap C = A \cap (B \cap C)$ (proprietà associativa dell'intersezione);
- $A \cap B = B \cap A$ (proprietà commutativa dell'intersezione).

Va però segnalato che un'analogia tra le operazioni aritmetiche e insiemistiche deve tenere presente alcune differenze. Ad esempio, ricordiamo le due seguenti proprietà che coinvolgono sia l'unione che l'intersezione:

$$\begin{aligned} A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \end{aligned}$$

Va però notato che soltanto la seconda ha un analogo in ambito aritmetico (detta proprietà *distributiva*)

$$n(m + k) = nm + nk$$

mentre non esiste una proprietà analoga alla prima. Infatti, $n + mk$ non è (in generale) uguale a $(n + m)(n + k)$.

Ancora a proposito di analogie, possiamo citare l'esistenza di elementi neutri in tutti i casi:

$$n + 0 = n$$

$$n1 = n$$

e così

$$A \cup \emptyset = A$$

$$A \cap X = A$$

Anche la funzione di annullatore dello 0 è riproducibile con l'insieme \emptyset :

$$n0 = 0$$

$$A \cap \emptyset = \emptyset$$

Tuttavia, nel caso dei sottoinsiemi di X si ha anche la formula simmetrica (duale)

$$A \cup X = X$$

che non corrisponde a nulla in \mathbb{N} .

Continuando con le diversità, possiamo notare che nell'insieme delle parti di X valgono equazioni del tipo:

$$A \cup A = A \text{ (idempotenza dell'unione)}$$
$$A \cap A = A \text{ (idempotenza dell'intersezione)}$$

ed esiste, per ogni A , un altro sottoinsieme di X , il suo complemento \bar{A} , tale che

$$A \cup \bar{A} = X \qquad \qquad A \cap \bar{A} = \emptyset$$

D'altra parte, \mathbb{N} risulta un insieme totalmente ordinato rispetto alla relazione \leq definita da:

$n \leq m$ se e soltanto se esiste k tale che $n +$

$$k = m$$

Invece, $\wp(X)$ risulta un insieme parzialmente ordinato rispetto alla relazione di inclusione \subseteq .

6.2 Algebre di Boole

Ai filosofi ed ai matematici che si proponevano di ridurre il calcolo delle classi al calcolo dei numeri (il più noto tra tutti è G. Leibniz), le ultime proprietà che abbiamo osservato nel precedente paragrafo sembrarono bizzarre e passò molto tempo finché vennero accettate come proprietà ammissibili per delle operazioni. Tuttavia, l'analogia tra il calcolo delle classi ed il calcolo delle proprietà con i connettivi sintattici, che abbiamo osservato più volte nei precedenti capitoli, spingeva ad uno

studio astratto delle operazioni di unione, intersezione, complemento, ecc. Si arrivò così all'introduzione dell'algebra di Boole (dal nome di George Boole) come caratterizzazione della struttura algebrica rilevata nell'insieme delle parti di un insieme X .

Definizione 6.1. Si dice algebra di Boole una sestupla $(B, \sqcap, \sqcup, {}^c, \perp, \top)$, dove

$$\sqcup : B \times B \rightarrow B \quad \sqcap : B \times B \rightarrow B \quad {}^c : B \rightarrow B \quad \perp \in B \quad \top \in B$$

tali che, per ogni $a, b, c \in B$:

$$(\textbf{Comm}) \quad a \sqcap b = b \sqcap a \text{ e } a \sqcup b = b \sqcup a;$$

$$(\textbf{Assoc}) \quad (a \sqcap b) \sqcap c = a \sqcap (b \sqcap c) \text{ e } (a \sqcup b) \sqcup c = a \sqcup (b \sqcup c);$$

$$(\textbf{Distr}) \quad a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c) \text{ e } a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c)$$

$$(\textbf{Idemp}) \quad a \sqcap a = a \text{ e } a \sqcup a = a;$$

$$(\textbf{Ident}) \quad a \sqcup \perp = a \text{ e } a \sqcap \top = a;$$

$$(\textbf{Compl}) \quad a \sqcap a^c = \perp \text{ e } a \sqcup a^c = \top.$$

L'elemento $a^c \in B$ è di solito detto *complemento*. Si verifichi che, per ogni insieme X , $(\wp(X), \cap, \cup, \neg, \emptyset, X)$ è un'algebra di

Boole.

Proposizione 6.1. *Sia $(B, \sqcap, \sqcup, {}^c, \perp, \top)$ un’algebra di Boole; allora è possibile definire su B una relazione d’ordine \sqsubseteq in modo che \sqcap dia il massimo comune minorante e \sqcup il minimo comune maggiorante.*

Dimostrazione. Si ponga $a \sqsubseteq b$ se e soltanto se $b \sqcup a = b$ (o, equivalentemente, $a \sqcap b = a$; nell’Esercizio 6.8 è chiesto di dimostrare che queste due possibili definizioni di \sqsubseteq sono equivalenti). La riflessività discende da (Idemp): per definizione, $a \sqsubseteq a$ se e soltanto se $a \sqcup a = a$, che è vero per via di (Idemp). L’antisimmetria discende da (Comm): se $a \sqcup b$ e $b \sqcup a$, per definizione si ha che $b \sqcup a = b$ e $a \sqcup b = a$; per (Comm), $a = a \sqcup b = b \sqcup a = b$. La transitività da (Assoc): se $a \sqsubseteq b$ e $b \sqsubseteq c$, per definizione si ha che $b \sqcup a = b$ e $c \sqcup b = c$, da cui $c \sqcup (b \sqcup a) = c$.

$= c$; per (Assoc), $c \sqcup (b \sqcup a) = (c \sqcup b) \sqcup a = c \sqcup a$.

Dobbiamo ora mostrare che $a \sqcap b$ è il massimo tra i minoranti di a e di b . Iniziamo col mostrare che è un minorante, cioè $a \sqcap b \sqsubseteq a$ (e, similmente, $a \sqcap b \sqsubseteq b$); dobbiamo cioè dimostrare che $a \sqcap (a \sqcap b) = a$. Per (Ident), (Distr) e (Comm), abbiamo che

$$a \sqcap (a \sqcap b) = (a \sqcap T) \sqcap (a \sqcap b) = a \sqcap (T \sqcap b) = a \sqcap (b \sqcap T) = a \sqcap T = a.$$

La massimalità di $a \sqcap b$ tra i minoranti di a e di b è provata dicendo che, se $x \sqsubseteq a$ e $x \sqsubseteq b$ (cioè, $a \sqcup x = a$ e $b \sqcup x = b$), allora $(a \sqcap x) \sqcap (b \sqcap x) = a \sqcap b$; da ciò otteniamo che $x \sqsubseteq a \sqcap b$, visto che $(a \sqcap x) \sqcap (b \sqcap x) = (x \sqcap a) \sqcap (x \sqcap b) = x \sqcap (a \sqcap b) = (a \sqcap b) \sqcap x$.

La prova che \sqcap da il minimo comun maggiorante è simile.

□

Si verifichi che, nell'algebra di Boole $(\wp(X), \cap, \cup, \neg, \emptyset, X)$, l'ordinamento definito

nella proposizione precedente coincide con \subseteq .

Proposizione 6.2. *Sia dato un insieme parzialmente ordinato (B, \sqsubseteq) nel quale*

- *per ogni coppia di elementi (a, b) esista un massimo comune minorante $a \sqcap b$ ed un minimo comune maggiorante $a \sqcup b$ (in questo caso (B, \sqsubseteq) viene detto reticolo),*
- *ogni elemento abbia un complemento, e*
- *valga (Distr).*

Sia c la funzione che associa ad ogni elemento il suo complemento e siano \perp e \top il massimo comune minorante ed il minimo comune maggiorante tra un qualsiasi $a \in B$ ed il suo complemento. Allora $(B, \sqcap, \sqcup^c, \perp, \top)$ è un'algebra di Boole.

Proposizione 6.3. *In un’algebra di Boole valgono le seguenti proprietà:*

1. $(a \sqcap b)^c = a^c \sqcup b^c$ e $(a \sqcup b)^c = a^c \sqcap b^c$ (*leggi di De Morgan*¹);
2. $(a^c)^c = a$;
3. $a^c \sqsubseteq b^c$ se e soltanto se $b \sqsubseteq a$.

Come si vede, l’operazione di complemento opera quella “simmetria” della struttura che abbiamo chiamato a suo tempo *dualità* (si veda il discorso a fine di Sezione 2.4). Infatti, l’operazione di complemento inverte l’ordinamento (si veda il terzo punto della proposizione precedente), inducendo la dualità tipica delle strutture d’ordine, questa volta, però, all’interno di una singola algebra di Boole.

Tra le algebre di Boole, riveste particolare importanza l’algebra **2**, costruita sull’insieme $B = \{0, 1\}$ con $\perp = 0$

$\sqsubseteq 1 = \top$. Diamo innanzitutto le tavole per le operazioni in $\mathbf{2}$, indicando \sqcup e \sqcap con $+$ e \cdot rispettivamente:

| | | + | | | • |
|---|---|---|---|---|---|
| | | | 0 | 1 | |
| 0 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 1 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 | 0 | 0 |
| | 1 | 1 | 1 | 1 | 1 |

Naturalmente i due elementi, 0 ed 1 , sono l'uno complementare dell'altro.

L'insieme delle parti di un insieme X è in corrispondenza biunivoca con le funzioni da X in $\mathbf{2}$, perchè ad ogni sottoinsieme di X possiamo associare la sua *funzione caratteristica*: dato un sottoinsieme A di X , possiamo definire la funzione $\chi_A: X \rightarrow \mathbf{2}$ che associa 1 agli elementi di X che sono in A e 0 agli altri. Viceversa, data una funzione $f: X \rightarrow \mathbf{2}$, possiamo considerare il sottoinsieme di X costituito dalle controimmagini di 1 . Nel caso di insiemi finiti (ma la cosa si può

estendere anche al caso generale), ciò significa che ad ogni sottoinsieme di un X con n elementi viene associata una n -upla di 0 e di 1. L'unione, l'intersezione ed il complementare possono allora essere calcolati punto per punto usando le tavole delle operazioni di **2** come nel seguente esempio.

Esempio 6.1. Siano $X = \{a, b, c, d, e\}$, $A = \{a, c, d\}$, $B = \{b, d\}$, $\chi_A =$

$$\begin{aligned}\chi_{A \cup B} &= (1+0, 0+1, 1+0, 1+1, 0+0) = (1, 1, 1, 1, 0) \\ \chi_{A \cap B} &= (1 \bullet 0, 0 \bullet 1, 1 \bullet 0, 1 \bullet 1, 0 \bullet 0) = (0, 0, 0, 1, 0) \\ \chi_{\bar{A}} &= (1^c, 0^c, 1^c, 1^c, 0^c) = (0, 1, 0, 0, 1)\end{aligned}$$

Averemo quindi che $A \cup B = \{a, b, c, d\}$, $A \cap B = \{d\}$ e $\bar{A} = \{b, e\}$. Si osservi che le somme ed i prodotti sono quelli delle algebre di Boole e, pertanto, operano cifra a cifra; quindi, nessun tipo di riporto o nozioni collegate sono da assumere.

Lo stesso discorso si può estendere facilmente al caso numerabile. Un sottoinsieme dei numeri naturali corrisponde ad una successione numerabile di 0 ed 1 che determina la sua funzione caratteristica. Come già osservato (si veda il Corollario 4.1), in questo modo un sottoinsieme dei numeri naturali corrisponde ad un numero reale in $(0, 1)$ scritto in base 2.

Una delle più importanti e classiche applicazioni della teoria delle algebre di Boole riguarda la costruzione e la semplificazione di *circuiti*. Naturalmente si tratta di circuiti in senso astratto perché quanto diremo sarà altrettanto bene riferibile sia a circuiti elettrici (o elettronici), che a circuiti idraulici o di traffico, ecc.. La caratteristica peculiare di questi circuiti dovrà essere la presenza di interruttori a due soli valori: “aperto” e

“chiuso”. Ciò significa che un circuito è una funzione $\mathbf{2}^n \rightarrow \mathbf{2}$. Infatti il passaggio o meno del “flusso di corrente” attraverso il circuito sarà determinato dai valori assunti dagli interruttori. La “somma” ed il “prodotto” di interruttori sarà rispettivamente la loro messa in parallelo o in serie: nel primo caso, la corrente passerà se passa per almeno uno dei due; nell’altro caso, la corrente passerà se passa per ambedue. Detto in altre parole, la funzione “somma” assume valore 1 se almeno uno dei due addendi assume valore 1, mentre la funzione “prodotto” assume valore 1 se ambedue i fattori assumono valore 1.

Abbiamo già accennato alla relazione tra le operazioni in $\wp(X)$ ed i connettivi logici dei quali ci occuperemo nel prossimo capitolo; a questo punto è chiaro che questi ultimi potranno essere

caratterizzati anche da circuiti elementari che prendono il nome di *porte logiche*.

Esercizi svolti

Esercizio 6.1. Si dimostri che, in un'algebra di Boole, valgono le seguenti leggi:

$$(i) a \sqcup \top = \top$$

$$(ii) \perp^c = \top$$

$$(iii) \perp \sqsubseteq a \sqsubseteq \top, \text{ per ogni } a$$

Soluzione:

(i) Per (Compl), si ha che $\top = a \sqcup a^c$, da cui $a \sqcup \top = a \sqcup (a \sqcup a^c)$; grazie a (Assoc), (Idemp) e (Compl), abbiamo che $a \sqcup (a \sqcup a^c) = (a \sqcup a) \sqcup a^c = a \sqcup a^c = \top$, come desiderato.

(ii) Per (Compl), si ha che $\top = \perp \sqcup \perp^c$. Poiché \perp è il minimo, deve essere che

il minim comune maggiorante tra \perp e qualsiasi altro elemento a è a stesso; pertanto, $\perp \sqcup \perp^c = \perp^c$, come richiesto.

(iii) Per il punto (i) di questo esercizio, abbiamo che $a_{a \sqcup \top} = \top$; per la definizione di \sqsubseteq data nelle prova della Proposizione 6.1, questo vuol dire che $a \sqsubseteq \top$. Per il principio di dualità, la (i) implica che $a \sqcap \perp = \perp$; di nuovo per la definizione di \sqsubseteq data nelle prova della Proposizione 6.1, questo equivale a $\perp \sqsubseteq a$.

Esercizio 6.2. *Si provi che in ogni algebra di Boole le seguenti condizioni sono equivalenti:*

$$(i) a \sqcap b^c = \perp \quad (ii) a \sqcup b = b \quad (iii) a^c \sqcup b = \top$$

Soluzione: Dire che due condizioni A e B sono equivalenti vuol dire che si riesce a dimostrare B assumendo A come ipotesi e

viceversa (più sinteticamente questo si esprime dicendo che A implica B e viceversa). Pertanto, basterà dimostrare che (i) implica (ii), che (ii) implica (iii) e che (iii) implica (i); tutte le altre implicazioni varranno di conseguenza.

(i) implica (ii): Per definizione di minimo comun maggiorante e di massimo, abbiamo che $a \sqcup b = (a \sqcup b) \sqcap \top$; per

(Compl), (Comm) e (Distr), abbiamo che

$$(a \sqcup b) \sqcap \top = (a \sqcup b) \sqcap (b \sqcup b^c) = (b \sqcup a) \sqcap (b \sqcup b^c) = b \sqcup (a \sqcap b^c)$$

. La condizione (i) e la definizione di minimo ci permettono di concludere che $b \sqcup (a \sqcap b^c) = b \perp = b$, come richiesto.

(ii) implica (iii): Per (ii), abbiamo che $ac \sqcup b = ac \sqcup (a \sqcup b)$; per (Assoc), (Comm) e (Compl), nonchè per definizione di massimo e minimo comun maggiorante, otteniamo che

$$a^c \sqcup (a \sqcup b) = (a^c \sqcup a) \sqcup b = (a \sqcup a^c) \sqcup b = \top \sqcup b = \top$$

(iii) implica (i): Per il punto 3 della Proposizione 6.3² applicato a (iii) e per il punto (ii) dell'Esercizio 6.1, abbiamo che $(a^c \sqcup b)^c = \top^c = \perp$; ora, per i punti 1 e 2 della Proposizione 6.3, si ha che $(a^c \sqcup b)^c = (a^c)^c \sqcap b^c = a \sqcap b^c$, e quindi (i).

Esercizio 6.3 (Unicità del complemento). *Si dimostri che, per ogni elemento a di un'algebra di Boole, esiste un unico elemento (che abbiamo indicato con ac) che soddisfi le condizioni descritte da (Compl).*

Soluzione: Che un tale elemento esista è garantito dalla definizione di algebra di Boole; dobbiamo quindi dimostrarne l'unicità. Le dimostrazioni di unicità tipicamente procedono assumendo l'esistenza di due elementi che godano

delle proprietà in esame e dimostrando che tali elementi devono essere uguali. Pertanto, assumeremo che esistano a' ed a'' tali che $a \sqcup a' = a \sqcup a'' = \top$ e $a \sqcap a' = a \sqcap a'' = \perp$; dobbiamo dimostrare che $a' = a''$.

Per definizione di minimo, per l'assunzione fatta su a'' e per (Distr) e (Comm), abbiamo che

$a' = a' \sqcup \perp = a' \sqcup (a \sqcap a'') = (a' \sqcup a) \sqcap (a' \sqcup a'') = (a \sqcup a') \sqcap (a' \sqcup a'')$. Per l'assunzione fatta su a'' , per (Comm) e per definizione di massimo, $(a \sqcup a') \sqcap (a' \sqcup a'') = \top \sqcap (a' \sqcup a'') = (a' \sqcup a'') \sqcap \top = a' \sqcup a''$; pertanto, abbiamo che $a' = a' \sqcup a''$. Con un analogo procedimento, possiamo dimostrare che $a'' = a'' \sqcup a'$. Poichè $a' \sqcup a'' = a'' \sqcup a'$, concludiamo che $a' = a''$, come desiderato.

Esercizio 6.4. *Si dimostri che in un'algebra di Boole, non vale la proprietà di cancellazione, né rispetto a \sqcap né rispetto*

$a \sqcup$.

Soluzione: Basta trovare tre elementi a , b e c (non necessariamente distinti) di un’algebra di Boole tali che $a \sqcup b = a \sqcup c$ ma per cui non vale $b = c$. Basti pensare all’algebra **2**, con $a = b = 1$ e $c = 0$: si ha che $1+1 = 1+0$ ma, ovviamente, 1 non è uguale a 0. Usando il principio di dualità, il lettore provi a trovare un controesempio alla proprietà di cancellazione per \sqcap .

Esercizio 6.5. *Sia $(B, \sqcap, \sqcup^c, \perp, \top)$ un’algebra di Boole, sia $a \in B$ e sia $B_a = \{b \in B : b \sqsubseteq a\}$. Si dimostri che $(B_a, \sqcap, \sqcup^c, *, \perp, a)$, dove $b* = b^c \sqcap a$, è ancora un’algebra di Boole.*

Soluzione: Anzitutto, osserviamo che \sqcap e \sqcup sono ancora funzioni, se ristrette a B_a . Infatti, per ogni $b, c \in B_a \subseteq B$, esistono (in B) gli elementi $b \sqcap c$ e $b \sqcup c$. Per

definizione, $b \sqcap c \sqsubseteq b \sqsubseteq a$ e quindi, per transitività, $b \sqcap c \in B_a$. Invece, essendo per definizione $b \sqsubseteq a$ e $c \sqsubseteq a$, si ha che a è un comune maggiorante sia di b che di c e quindi $b \sqcup c \sqsubseteq a$, essendo $b \sqcup c$ il minimo dei comuni maggioranti; quindi $b \sqcup c \in B_a$. Chiaramente, anche se ristrette a B_a , le funzioni \sqcap e \sqcup godono delle proprietà (Comm), (Assoc), (Idemp) e (Distr) poichè per ipotesi godono di tali proprietà in B .

Anche la proprietà (Ident) è facile da dimostrare: per ogni $b \in B_a$, si ha che $\perp \sqsubseteq b$, poichè ciò valeva in B ; pertanto, per definizione di \sqsubseteq , si ha che $b \sqcup \perp = b$. Inoltre, per costruzione di B_a , si ha che $b \sqsubseteq a$; quindi, grazie all'[Esercizio 6.8](#), $a \sqcap b = b$ che, per (Comm), implica $b \sqcap a = b$.

Dobbiamo ora mostrare che $*$ in B_a è una funzione e soddisfa la proprietà (Compl). Chiaramente, per ogni $b \in B_a$,

esiste ed è unico l'elemento b^* : infatti, per ogni $b \in B$, esiste ed è unico $b^c \sqcap a$, che inoltre appartiene a B_a essendo $b^c \sqcap a \sqsubseteq a$.

Sia ora $b \in B_a$; per le proprietà (Assoc), (Compl), (Comm) e (Ident) in B , si ha che

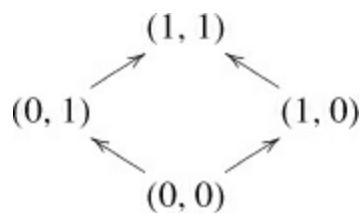
$$b \sqcap b^* = b \sqcap (b^c \sqcap a) = (b \sqcap b^c) \sqcap a = \perp \sqcap a = a \sqcap \perp = \perp.$$

Inoltre, per le proprietà (Distr), (Compl), (Ident) e (Comm) in B , nonchè per

$$b \sqcup b^* = b \sqcup (b^c \sqcap a) = (b \sqcup b^c) \sqcap (b \sqcup a) = \top \sqcap (b \sqcup a) = b \sqcup a = a \sqcup b = a.$$

Esercizio 6.6. *Si definisca un'algebra di Boole con quattro elementi.*

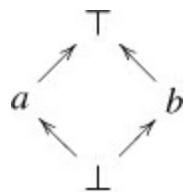
Soluzione: Si considerino gli elementi di $\{0, 1\} \times \{0, 1\}$ ordinati per componenti:



Non è difficile mostrare che tale insieme ordinato soddisfa le tre proprietà

specificate nell'ipotesi della Proposizione 6.2. Usando tale risultato, possiamo concludere.

In generale, possiamo prendere $B = \{\perp, a, b, \top\}$ e definire:



Esercizi da svolgere

Esercizio 6.7. Si dimostrino i tre punti della Proposizione 6.3.

Esercizio 6.8. Si provi che $a \sqcup b = a$ se e solo se $a \sqcap b = b$.

Esercizio 6.9 (Unicità dell'identità). Si dimostri che in un'algebra di Boole \top e sono gli unici elementi che soddisfano

(Ident).

Esercizio 6.10. Sia $(B, \sqcap, \sqcup^c, \perp, \top)$ un’algebra di Boole. Si dimostri che $(B, \sqcup, \sqcap^c, \top, \perp)$ è ancora un’algebra di Boole, chiamata algebra duale.

Esercizio 6.11. Si definisca un’algebra di Boole con n elementi, dove $n \in \mathbb{N}$.

Esercizio 6.12. Si consideri $F = \{X \subseteq \mathbb{N} : X \text{ è un insieme finito}\}$ e sia $Cof F = \{\overline{X} : X \in F\}$. La sestupla $(F \cup Cof, \sqcap, \sqcup, \neg, \emptyset, \mathbb{N})$ è un’algebra di Boole? Se sì, che cardinalità ha? Cosa cambierebbe se $F = \{X \subseteq \mathbb{N}\}$?

Esercizio 6.13. Siano $(B_1, \sqcap_1, \sqcup_1, {}^{c_1}, \perp_1, \top_1)$ e $(B_2, \sqcap_2, \sqcup_2, {}^{c_2}, \perp_2, \top_2)$ due algebre di Boole, e sia $f: B_1 \rightarrow B_2$ una funzione tale che, per ogni $a, b \in B_1$, si ha che

$f(a \sqcap_1 b) = f(a) \sqcap_2 f(b)$ e $f(a \sqcup_1 b) = f(a) \sqcup_2 f(b)$. Si dimostri che le seguenti affermazioni

sono equivalenti:

(a) $F \cup Cof, \cap, \cup, \neg, \emptyset, \mathbb{N}$

(b) $f(a^c_1) = (f(a))^c_2$, per ogni $a \in B_1$.

¹ L'opera di Augustus De Morgan (1806-1871) fornì un'interpretazione simbolica del sillogismo aristotelico efficace e semplice, nella quale, correttamente, l'eliminazione del termine medio appare come il momento essenziale dell'intero procedimento di inferenza. Alcune intuizioni di De Morgan, inoltre, anticiparono importanti ricerche del XIX secolo: dunque, l'importanza storica del sistema di De Morgan può essere considerata in relazione allo sviluppo dell'algebra delle relazioni di Peirce, nonchè per aver fornito a Boole uno spunto importante collegato ai suoi tentativi di generalizzazione della sillogistica mediante il calcolo delle classi.

² Si noti che \sqsubseteq è una relazione d'ordine e quindi, grazie alla proprietà antisimmetrica, il punto 3 della Proposizione 6.3 ci permette di concludere che $a = b$ se e solo se $a^c = b^c$.

Parte B

Introduzione alla logica matematica

La formalizzazione matematica

Nel tentativo di ricerca della verità e della correttezza del ragionamento, la civiltà greca si dedicò all'analisi del linguaggio ed al tentativo di applicare al ragionamento corrente quelle regole rigorose che si erano mostrate così proficue nel costruire il monumento matematico rappresentato dagli *Elementi* di Euclide. Il dibattito tra Sofisti e Filosofi non era in fondo altro che un dibattito sulle possibilità del linguaggio, che gli uni ritenevano intrinsecamente ambiguo mentre gli altri

si sforzavano di rendere rigoroso fino al punto di diventare inattaccabile. Questo tentativo, per certi versi anche un po' ingenuo se portato avanti in maniera troppo assoluta, si è protratto attraverso il Medioevo fino a Leibniz, che vi dedicò buona parte della sua ricerca, rimasta poi per lungo tempo sconosciuta. Dal canto suo, la stessa matematica stava diventando sempre più formale ed astratta nel tentativo di rendere il controllo delle sue costruzioni e dei suoi metodi sempre più rigoroso ed efficiente. Dal XIX secolo in poi, assistiamo ad un grande sviluppo di questo aspetto di formalizzazione di diversi linguaggi, tra i quali abbiamo già visto quello della teoria degli insiemi e delle algebre di Boole.

Più in generale, lo sviluppo di una qualsiasi nozione di *calcolo* comporta l'astrazione dalla natura dei termini sui

quali il calcolo stesso viene effettuato e la loro manipolazione attraverso le rigide regole che governano quel particolare calcolo.

Come esempio si pensi ai giochi: nel gioco degli scacchi un *cavallo* non ha niente a che fare con la razza equina, ma è semplicemente quel pezzo (termine) soggetto ad una particolare regola di movimento sulla scacchiera che lo distingue da tutti gli altri pezzi. Ancora più significativi per i nostri scopi sono i giochi di carte, dove la stessa carta è soggetta a regole diverse al cambiare del gioco e cambia quindi di valore (significato) di conseguenza.

Ciò accade anche nel linguaggio corrente, che è soggetto di volta in volta a diversi “giochi linguistici” (Wittgenstein). Si pensi al significato diverso che può assumere una stessa frase in un contesto

serio ed in uno scherzoso. Tanto più ciò accade quando si ha a che fare con linguaggi specifici. Ad esempio, il termine *vettore* viene usato per indicare cose diverse in matematica/fisica e nel linguaggio di una agenzia di viaggi.

Le scienze ormai sono ritenute tali soltanto se suscettibili in qualche modo di una matematizzazione; questo aspetto risulta assolutamente imprescindibile quando si parla di informatica. Qui, infatti, ritroviamo i due oggetti principali della formalizzazione: il *calcolo* ed il *linguaggio*.

Il ragionamento può presentarsi sotto forma di calcolo in diversi modi.

Esempio 7.1. *Si considerino i due seguenti ragionamenti, nei quali il terzo enunciato viene pensato come conseguenza dei primi due.*

1. • *Se piove e le strade sono bagnate allora la circolazione diventa pericolosa;*
 - *se piove allora le strade sono bagnate;*
 - *se piove allora la circolazione diventa pericolosa.*

2. • *Se mi iscrivo all'università e studio allora supererò gli esami;*

- *se mi iscrivo all'università allora studio;*
- *se mi iscrivo all'università allora supererò gli esami.*

Si può vedere facilmente che i due ragionamenti sono formalmente uguali:

3. • *Se (A e B) allora C;*
 - *se A allora B;*
 - *se A allora C.*

Altrettanto facilmente si vede che non li valuteremmo esattamente nello stesso modo nel nostro modo di ragionare corrente.

Nel momento in cui riduciamo a calcolo il ragionamento, dobbiamo certamente sacrificare alcune “sfumature”, ma acquisiamo garanzie nel funzionamento di certi schemi.

Si tratta allora di stabilire opportuni calcoli, cioè dei sistemi di regole a seconda dei casi che vogliamo trattare. Nei giochi sopra menzionati, le regole sono puramente di fantasia o, almeno, fissate in modo da rendere il gioco divertente e con una ragionevole probabilità di vittoria. Nel nostro caso non siamo così liberi: occorre che le regole che imponiamo rispecchino il più fedelmente possibile quanto noi intendiamo per “ragionamento corretto”. In altre parole, vogliamo che ogni volta

che sostituiamo A, B e C con enunciati del nostro universo di discorso le cose vadano ragionevolmente bene. In maniera più rigorosa, diciamo che stiamo costruendo una *sintassi*, cioè un sistema di regole per la manipolazione dei termini, che deve corrispondere ad una data *semantica*, cioè a quello che noi intendiamo, ogniqualvolta *interpretiamo* la sintassi. Nell'esempio precedente, i primi due ragionamenti sono interpretazioni dello schema sintattico proposto come terzo. Dovrà accadere che, se una volta dato un sistema di regole per le quali lo schema sintattico viene accettato, allora tutte le sue interpretazioni dovranno essere vere (*correttezza del sistemacorrettezza*); viceversa, se tutte le interpretazioni di uno stesso schema sintattico risultano vere, il sistema dovrà essere in grado di accettare lo schema stesso (*completezza del sistemacompletezza*). Mentre la prima

proprietà è assolutamente imprescindibile, non saremo sempre in grado di soddisfare la seconda.

Quali calcoli ci interessano? Nell'esempio sopra proposto abbiamo visto un “gioco” che riguardava i *connettivi*, cioè quegli elementi del discorso che connettono diversi enunciati, ad esempio: “e”, “o”, “se... allora”, ecc. In questo caso gli enunciati sono pensati come “scatole chiuse”, ovvero pedine tutte dello stesso tipo, distinguibili soltanto dal nome.

Potremmo invece pensare di “aprire le scatole” ed andare a vedere se possiamo fare dei ragionamenti in base a come sono fatte dentro, anche senza fare appello alla semantica.

Esempio 7.2. *Si considerino i due ragionamenti nei quali il terzo enunciato viene pensato come conseguenza dei primi due:*

1. • *Tutti gli uomini sono mortali;*
• *Socrate è un uomo;*
• *Socrate è mortale.*

2. • *Tutti i gatti hanno cinque zampe;*
• *Fuffi è un gatto;*
• *Fuffi ha cinque zampe.*

I due ragionamenti sono formalmente uguali:

- *Tutti gli A sono B;*
- *a è A;*
- *a è B.*

Tuttavia, come prima, non li valuteremmo nello stesso modo nel nostro modo di ragionare corrente.

Si noti che, diversamente dall'[Esempio 7.1](#), in questo caso abbiamo a che fare ogni volta con un singolo enunciato (cioè, non

ci sono connettivi) ed il ragionamento si basa su come tale enunciato è fatto all'interno, in sostanza sui *quantificatori*, i *termini individuali* ed i *predicati* che vengono coinvolti. Ci troviamo quindi di fronte a due tipi di calcolo sostanzialmente diversi (ne esistono anche altri, ma che esulano dai limiti di questo volume): quello degli *enunciati-connettivi* e quello dei *predicati-quantificatori*. Storicamente e razionalmente, il secondo precede il primo, ma per motivi didattici preferiamo seguire l'ordine inverso, essendo il primo di gran lunga più semplice.

Calcolo degli enunciati

8.1 Enunciati, connettivi e valori di verità

8.1.1 Verità

“I matematici sono stati sempre persuasi di dimostrare delle *verità* o delle *proposizioni vere*. Evidentemente tale convinzione non può essere che di ordine sentimentale o metafisico: non la si può certo giustificare ponendosi sul terreno della matematica”.

Nicolas Bourbaki [10]

La frase di Bourbaki che abbiamo scelto per introdurre la prima sezione dedicata alla logica potrà stupire il lettore: la logica appare infatti strettamente collegata alla nozione di verità e la posizione espressa dal grande matematico policefalo può apparire eccessivamente prudente.

Ricordando le radici del pensiero matematico e logico, troviamo che:

“Se andiamo a rileggere le dimostrazioni della geometria greca per considerarle non rispetto alla loro impostazione tecnica, ma per il ruolo che hanno nei confronti della tesi cui arrivano, ci troviamo sostanzialmente di fronte a tre situazioni principali che la dicono lunga sul ruolo assegnato alla dimostrazione: dimostrazioni di qualcosa che è vero, ma sconosciuto e

non immediato (...); dimostrazioni di qualcosa che è evidente, ad esempio che gli angoli alla base di un triangolo isoscele sono uguali; dimostrazioni di qualcosa che è contro l'intuizione.”

([6], Introduzione di A. Labella)

Dunque le dimostrazioni possono riguardare fatti più o meno evidenti o plausibili, ma riguardano comunque fatti *veri*. La logica di cui ci occupiamo è *bivalente*, cioè prevede che le espressioni assumano uno ed uno solo tra i due *valori di verità*: “vero”, V, o “falso”, F.

Intuitivamente, con l’attribuzione di uno di questi due valori si indica che la verità (o la falsità) dell’espressione in questione può essere stabilita senza dubbio, che è evidente e oggettiva.

La logica contemporanea ha però

evidenziato che il concetto di verità è delicato. Per introdurre la questione, proponiamo alcuni brani tratti da *Verità e dimostrazione* di Alfred Tarski [52] (riprenderemo alcune considerazioni da [3]) in cui si osserva che l'interpretazione di tale concetto ha radici antiche e può basarsi su considerazioni filosofiche come quelle espresse nel seguente passo, tratto dalla *Metafisica* di Aristotele:

“Dire di ciò che è che non è, o di ciò che non è che è, è falso, mentre dire di ciò che è che è o di ciò che non è che non è, è vero.”

(A. Tarski [52])

Se considerassimo queste affermazioni alla stregua di una “definizione”, dovremmo osservare che la formulazione sarebbe insufficiente dal punto di vista

formale: non è infatti abbastanza generale, in quanto è riferita soltanto a proposizioni che affermano qualche cosa (“che è” o “che non è”) di un soggetto, e sarebbe talvolta difficile far rientrare una proposizione qualsiasi in questo schema senza modificarne il senso.

[3] si propone di ottenere una soddisfacente spiegazione del concetto classico di verità, superando l'originale formulazione aristotelica ma conservando di essa gli intenti principali. Per fare ciò è innanzitutto indispensabile riferirsi ad un linguaggio: considereremo la lingua italiana. Seguiamo ancora Tarski ed esaminiamo la proposizione:

“La neve è bianca”

Che cosa intendiamo dire quando affermiamo che essa è vera o che è falsa?

Accettando l'impostazione di Aristotele potremmo scrivere [52]:

“La neve è bianca” è vera se e solo se la neve è bianca

“La neve è bianca” è falsa se e solo se la neve non è bianca

Queste frasi illustrano il significato dei termini “vero” e “falso” quando tali termini sono riferiti alla proposizione “la neve è bianca” e possono dunque essere considerate definizioni (parziali) dei termini “vero” e “falso”, cioè definizioni di questi termini in relazione alla particolare proposizione “La neve è bianca”.

Generalizzando quanto affermato per una singola proposizione, otteniamo (sempre seguendo [52]):

“*p*” è vera se e solo se *p*

(*)]

Nella proposizione che verrà sostituita a *p*, però, non dovrà comparire la parola “vero”, altrimenti la (*) verrebbe a costituire, ovviamente, un circolo vizioso e non sarebbe accettabile come definizione (parziale) di verità.

Quando avremo precisato un’equivalenza della forma (*) nella quale “*p*” sia sostituita da un’arbitraria proposizione italiana, potremo dire che l’uso del termine “vero” in riferimento alle proposizioni italiane è conforme al concetto classico di verità. Potremo allora affermare, nelle parole di Tarski, che “l’uso del termine “vero” è adeguato” [52].

Ma è possibile realizzare tutto ciò? È cioè possibile fissare un uso adeguato del termine “vero” per le proposizioni scritte nel linguaggio scelto (nella lingua italiana)?

A questo punto è necessario precisare l’ambito nel quale vogliamo definire il concetto di verità: il procedimento sopra descritto non sarebbe ad esempio applicabile considerando l’intera lingua italiana. Innanzitutto l’insieme delle proposizioni italiane è (potenzialmente) infinito; inoltre la parola “vero” compare nella lingua italiana e ciò impedisce di applicare il procedimento. Ma si presentano anche altri e ben più gravi problemi: se immaginassimo la possibilità di determinare un uso adeguato del termine “vero” con riferimento a proposizioni italiane del tutto arbitrarie, cadremmo inevitabilmente in una contraddizione: ci ritroveremmo infatti di fronte alla preoccupante possibilità di incontrare l’antinomia del mentitore (si veda la Sezione 4.4). Seguiamo l’esposizione di Tarski:

“Il linguaggio comune è universale, nè deve essere altrimenti, giacchè ci si aspetta che esso fornisca i mezzi adeguati a esprimere ogni cosa che possa essere espressa (...) Possiamo perfino costruire nel linguaggio ciò che talvolta viene detta una proposizione autologa, cioè una proposizione S che esprime il fatto che S stessa è vera o che è falsa. Se S esprime la propria falsità, si può dimostrare con un semplice ragionamento che S è contemporaneamente vera e falsa, e così ci ritroviamo di fronte l’antinomia.”

(A. Tarski [52])

Un grave problema è dunque determinato dalla potenza del linguaggio in cui scegliamo di operare; ma linguaggi universali non sono, in generale, assolutamente necessari per gli scopi della

ricerca scientifica. È allora possibile dare una definizione del concetto di verità per linguaggi semanticamente limitati?

Tarski risponde affermativamente, ma precisa alcune condizioni: è necessario che il vocabolario del linguaggio in questione sia completamente determinato e che siano formulate esplicitamente delle precise regole sintattiche sulle quali basare la formazione delle proposizioni. Tali regole devono essere formali, dunque riferite esclusivamente alla forma esteriore delle espressioni. I linguaggi che soddisfano a queste condizioni sono detti *formalizzati*.

Si noti inoltre che il linguaggio che è l'oggetto dello studio (per il quale dunque si vuole costruire la definizione di verità) non coincide con il linguaggio nel quale la definizione viene formulata; quest'ultimo si dice *metalinguaggio*, mentre il primo è

denominato *linguaggio oggetto*. Il metalinguaggio deve contenere come parte propria il linguaggio oggetto; deve inoltre contenere nomi per le espressioni del linguaggio oggetto e altri termini necessari allo studio del linguaggio oggetto. Sottolineiamo che nel procedimento di definizione del concetto di verità i termini semantici (ovvero quelli che collegano le proposizioni del linguaggio oggetto e gli oggetti a cui esse sono riferite) devono poter essere introdotti nel metalinguaggio mediante opportune definizioni. Tutto ciò conferma che il metalinguaggio deve essere più ricco del corrispondente linguaggio oggetto.

Considerate queste precisazioni è possibile concludere:

“Se tutte le precedenti condizioni sono soddisfatte, la costruzione della

desiderata definizione di verità non presenta difficoltà essenziali.

Tecnicamente, tuttavia, essa è troppo complicata per essere esposta qui in dettaglio. Per ogni data proposizione del linguaggio oggetto si può facilmente formulare la corrispondente definizione parziale della forma (*).”

(Tarski, 1969)

Si presenta infine un’ulteriore difficoltà: l’insieme costituito da tutte le proposizioni del linguaggio oggetto è infinito, mentre ogni proposizione del metalinguaggio è una sequenza finita di segni; pertanto non si può pensare di dare la desiderata definizione generale mediante un puro e semplice accostamento di tutte le (infinite) definizioni parziali. Eppure, conclude Tarski nel lavoro citato, la nostra definizione generale non è poi molto

diversa, almeno intuitivamente, da quell'accostamento:

“Molto approssimativamente, si procede come segue. Dapprima si considerano le proposizioni più semplici, che non contengono altre proposizioni come parti; per queste proposizioni si trova il modo di definire la verità direttamente (usando la stessa idea che conduce alle definizioni parziali). Poi, mediante l'uso delle regole sintattiche che riguardano la formazione di proposizioni più complicate a partire da quelle più semplici, si estende la definizione a proposizioni composte arbitrarie; si applica qui il metodo conosciuto in matematica come definizione per ricorsione”.

(Tarski, 1969)

8.1.2 Enunciati

Il paragrafo precedente mostra che il concetto di verità (o di falsità) di un'affermazione è certamente delicato e complesso. Un'impostazione rigorosa della logica matematica potrebbe allora concentrarsi innanzitutto sulle espressioni che possono essere scritte utilizzando (sintatticamente) un assegnato alfabeto e soltanto successivamente occuparsi della semantica di tali espressioni, ovvero dell'attribuzione di un significato e dei conseguenti valori di verità ad esse.

Dunque anche l'introduzione del concetto di enunciato, che come vedremo è strettamente collegata all'attribuzione dei valori di verità ad un'espressione, potrebbe essere rimandata. Tuttavia, didatticamente è utile anticipare sin d'ora che diremo *enunciato* o *proposizione* un'affermazione che assume uno ed un

solo valore di verità, vero oppure falso. E tale caratteristica è tutt’altro che banale: infatti non tutte le affermazioni assumono incontestabilmente uno ed un solo valore di verità.

Esempio 8.1. (*Controesempio*).

L’affermazione “Esiste almeno un numero reale tale che il suo quadrato sia il reale z” non è un enunciato: esso dipende dal particolare z che sarà considerato; la scelta di un valore z negativo o non negativo comporta un valore di verità rispettivamente falso o vero per l’affermazione data.

L’affermazione “Tutti i naturali pari maggiori di 2 sono somme di due numeri primi” può essere considerato un vero e proprio enunciato? Si tratta infatti della celebre congettura di Goldbach, un problema che abbiamo presentato nella

Sezione 5.3.3; com’è noto, nessun matematico sino ad oggi è stato in grado di dimostrare o smentire tale affermazione. In altri termini, non sappiamo se la frase sopra riportata sia vera o sia falsa; a rigore, non potremmo neppure essere sicuri che sia possibile stabilire la sua verità o la sua falsità !

Alcuni enunciati sono costituiti da una sola affermazione (come “La neve è bianca” citato da Tarski) e sono detti *enunciati atomici*. Sottolineiamo sin d’ora che in questo primo capitolo dedicato alla logica degli enunciati prescinderemo dalla “struttura interna” degli enunciati in questione: in effetti sarebbe importante esaminare il tipo di affermazione di volta in volta considerata, che spesso viene riferita ad un soggetto *variabile* (cioè essa può essere riferita ad un singolo soggetto ma anche ad un insieme di soggetti).

Questa nostra scelta è esclusivamente didattica e provvisoria: verrà superata quando passeremo alla considerazione della logica dei predicati.

Pur senza esaminare in questa fase la struttura degli enunciati, gli enunciati atomici non saranno gli unici che prenderemo in considerazione. Enunciati più complessi sono costituiti da più affermazioni, collegate da opportune parole (dette *connettivi*) come *o*, *e*, *se...* *allora...*, *se e solo se*. I connettivi collegano gli enunciati senza riguardo al significato che quelli possono assumere: l'unica caratteristica che viene indicata nella loro definizione è quale valore di verità abbia l'enunciato composto a partire soltanto dai valori di verità assegnati agli enunciati componenti.

Esempio 8.2. (*Controesempio*).
Intuitivamente:

- A = “il numero otto è rappresentato ad una sola cifra”
- B = “un triangolo ha tre lati”

sono enunciati veri; però il buon senso ci porterebbe a dire che

- “è inevitabile che A ” è falso (possiamo infatti rappresentare otto in base 2, ottenendo 1000); invece
- “è inevitabile che B ” è vero.

Da ciò potremmo concludere che l’operatore “è inevitabile che” non agisce sugli enunciati come fanno i “connettivi” logici, cioè tenendo conto esclusivamente dei valori di verità.

Di conseguenza nel paragrafo seguente, dopo aver definito sintatticamente gli enunciati composti mediante connettivi, ne stabiliremo il “comportamento”

facendo riferimento al valore di verità che ci aspettiamo debba valere in base alla verità delle singole componenti. Come detto nel precedente capitolo, faremo riferimento alla semantica, che, in questo caso, è data semplicemente dai valori di verità. Si tratta naturalmente di una grossa semplificazione, ma non è irragionevole pensare che un “mondo esterno” sia descrivibile mediante la verità e la falsità di un insieme di enunciati. Ad esempio, una giornata può essere metereologicamente descritta da:

- “C’è il sole” è vero;
- “Tira vento” è falso;
- “La temperatura è inferiore a 20 gradi” è falso;
- “Si prevedono condizioni stabili” è vero;
- ecc.

Tenendo conto di questo tipo di

semantica, i connettivi, che sono, in fondo, operazioni tra enunciati, saranno definiti mediante le *tavole di verità*.

8.1.3 Conngettivi e valori di verità

I connettivi formalizzano alcune parole e sono indicati da opportuni simboli; riportiamo quelli maggiormente usati:

$\neg A$ che formalizza “non A”

$A \wedge B$ che formalizza “A e B”

$A \vee B$ che formalizza “A o B”

$A \rightarrow B$ che formalizza “se A allora B”

$A \leftrightarrow B$ che formalizza “se A allora B e se B allora A”

Talvolta “non” viene indicato come “operatore” e non come “connettivo” in quanto, a differenza degli altri connettivi, non collega due enunciati ma opera su di un solo enunciato.

Esempio 8.3. (*Controesempio*). E' utile osservare che l'indicazione dei connettivi mediante congiunzioni come "non", "e", "o" richiede una qualche prudenza. Ad esempio, la scrittura $A \vee B$, che come sopra detto formalizza "A o B", deve essere intesa in senso inclusivo ("o A o B o entrambi"), non in senso esclusivo ("oAoS ma non entrambi").

Un'effettiva definizione dei connettivi richiede la precisazione delle tavole di verità, di cui ci occuperemo nel seguito del paragrafo.

Grazie ai connettivi è possibile definire induttivamente l'insieme degli enunciati: utilizzeremo un alfabeto costituito da lettere maiuscole (con le quali rappresenteremo gli enunciati atomici), dai connettivi sopra introdotti e da un insieme finito di segni come virgole o parentesi.

Possiamo allora procedere per induzione ed affermare che:

- A, B, C, \dots sono enunciati (atomici);
- se X, Y sono enunciati, allora $\neg X, X \wedge Y, X \vee Y, X \rightarrow Y$ e $X \leftrightarrow Y$ sono enunciati.

Nella seguente tabella (anche detta *tavola di verità*) sono riassunte le definizioni dei connettivi:

| A | B | $\neg A$ | $A \wedge B$ | $A \vee B$ | $A \rightarrow B$ | $A \leftrightarrow B$ |
|-----|-----|----------|--------------|------------|-------------------|-----------------------|
| V | V | F | V | V | V | V |
| V | F | (F) | F | V | F | F |
| F | V | V | F | V | V | F |
| F | F | (V) | F | F | V | V |

Costruiamo ora le tavole di verità di alcuni enunciati composti (il metodo della costruzione della tavola di verità è stato introdotto da Ludwig Wittgenstein).

Esempio 8.4. *La tavola di verità*

*dell'enunciato composto
 $\neg((A \wedge B) \vee (A \wedge \neg B))$ è:*

| A | B | $A \wedge B$ | $\neg B$ | $A \wedge \neg B$ | $(A \wedge B) \vee (A \wedge \neg B)$ | $\neg((A \wedge B) \vee (A \wedge \neg B))$ |
|-----|-----|--------------|----------|-------------------|---------------------------------------|---|
| V | V | V | F | F | V | F |
| V | F | F | V | V | V | F |
| F | V | F | F | F | F | V |
| F | F | F | V | F | F | V |

Osserviamo che i valori di verità di $\neg[(A \wedge B) \vee (A \wedge \neg B)]$ corrispondono a quelli di $\neg A$.

Esempio 8.5. *La tavola di verità dell'enunciato composto $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$ è:*

| A | B | $A \rightarrow B$ | $\neg B$ | $\neg A$ | $\neg B \rightarrow \neg A$ | $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$ |
|-----|-----|-------------------|----------|----------|-----------------------------|---|
| V | V | V | F | F | V | V |
| V | F | F | V | F | F | V |
| F | V | V | F | V | V | V |
| F | F | V | V | V | V | V |

Osserviamo che il valore di verità è V per ogni scelta dei valori di verità di A e di B .

Prima di procedere, osserviamo che l'introduzione di cinque connettivi è sovrabbondante: ad esempio, sarebbe stato molto più sintetico introdurre solamente “ \neg ” (non) e “ \rightarrow ” (se... allora). Avremmo allora ricondotto gli altri connettivi a combinazioni di questi; si verifica infatti che (lasciamo al lettore la costruzione delle tavole di verità degli enunciati composti):

$A \wedge B$ ha la stessa tavola di verità di
 $\neg(A \rightarrow \neg B)$;

$A \vee B$ ha la stessa tavola di verità di
 $\neg A \rightarrow B$;

$A \leftrightarrow B$ ha la stessa tavola di verità di
 $\neg((A \rightarrow B) \rightarrow \neg(B \rightarrow A))$.

Esempio 8.6. *Non è difficile verificare (ed il lettore lo farà facilmente) che gli enunciati composti $\neg(A \vee B)$ e $\neg A \wedge \neg B$ hanno gli stessi valori di verità. Questa*

osservazione (legge di De Morgan) ha delle conseguenze interessanti: in particolare, riflettendo su di essa possiamo renderci conto che il corretto uso della simbologia comunemente usata in matematica presuppone un'effettiva conoscenza delle relazioni tra i connettivi logici.

Consideriamo ad esempio l'equazione $x^2 = 1$. Le sue soluzioni si trovano spesso espresse compattamente nella forma $x = \pm 1$, intendendo con ciò che la x può assumere sia il valore $+1$ che il valore -1 . Dunque, utilizzando i connettivi logici, la precedente scrittura può essere espressa, più correttamente, da $x = 1 \vee x = -1$.

Consideriamo ora la scrittura $x^2 \neq 1$, che porta alla $x \neq \pm 1$. In questo caso, al simbolo “ \pm ” non è direttamente legato un connettivo “ \vee ”; ovvero, la precedente

scrittura non deve essere tradotta in $x \neq 1 \vee x \neq -1$ in quanto questa richiederebbe il verificarsi di almeno una delle condizioni $x \neq 1$ e $x \neq -1$ (quindi alla x potrebbe essere sostituito... un qualsiasi numero reale!), mentre $x^2 \neq 1$ richiede il contemporaneo verificarsi di entrambe tali condizioni. Ricordiamo piuttosto che $x^2 \neq 1$ deve essere interpretata come la negazione di $x^2 = 1$; dunque essa corrisponde a $\neg(x = \pm 1)$ cioè $\neg((x = 1) \vee (x = -1))$ e infine $\neg(x = 1) \wedge \neg(x = -1)$.

Alla luce di quanto osservato, possiamo dire che i connettivi binari sono 16 in tutto (sono infatti quanti le funzioni da un insieme con quattro elementi ad uno con due), ma sono interdefinibili. In realtà, dal momento che di un connettivo interessa soltanto la tavola di verità, questa si può ottenere da una combinazione di altri connettivi. Ad esempio, l'insieme $\{\neg, \wedge,$

$\vee\}$ è sufficiente a definirli tutti, ma anche $\{\neg, \rightarrow\}$ lo è o, addirittura $\{\uparrow\}$, che corrisponde a “nè... nè” (esercizio). Tali insiemi si dicono *basi* di connettivi.

8.1.4 Interpretazioni, equivalenza logica e validità

Diremo *interpretazione* di un enunciato composto una funzione che assegna uno dei due valori di verità V o F a ciascun enunciato atomico componente e che quindi assegna un valore di verità all'enunciato composto sulla base delle tavole di verità.

Definizione 8.1. *Due enunciati A e B si dicono logicamente equivalenti (e scriveremo $A \equiv B$) se hanno lo stesso valore di verità per ogni interpretazione.*

Esempio 8.7. *Le seguenti sono*

equivalenze logiche:

$$A \equiv \neg \neg A$$

$$A \equiv A \wedge A$$

$$A \equiv A \vee A$$

$$A \wedge B \equiv B \wedge A$$

$$A \vee B \equiv B \vee A$$

$$A \leftrightarrow B \equiv B \leftrightarrow A$$

$$A \rightarrow B \equiv \neg B \rightarrow \neg A$$

$$A \wedge (B \wedge C) \equiv (A \wedge B) \wedge C$$

$$A \vee (B \vee C) \equiv (A \vee B) \vee C$$

$$A \leftrightarrow (B \leftrightarrow C) \equiv (A \leftrightarrow B) \leftrightarrow C$$

$$A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$$

$$A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$$

$$A \wedge (A \vee B) \equiv A$$

$$A \vee (A \wedge B) \equiv A$$

$$A \leftrightarrow B \equiv (A \rightarrow B) \wedge (B \rightarrow A)$$

$$A \rightarrow B \equiv \neg A \vee B$$

$$A \rightarrow B \equiv \neg (A \wedge \neg B)$$

$$A \wedge B \equiv \neg (\neg A \vee \neg B) \text{ (legge di De Morgan)}$$

Morgan)

$$A \vee B \equiv \neg(\neg A \wedge \neg B) \text{ (*legge di De Morgan*)}$$

$$A \wedge B \equiv \neg(A \rightarrow \neg B)$$

$$A \vee B \equiv \neg A \rightarrow B$$

$$A \rightarrow B \equiv A \leftrightarrow (A \wedge B)$$

$$A \rightarrow B \equiv B \leftrightarrow (A \vee B)$$

$$A \wedge B \equiv (A \leftrightarrow B) \leftrightarrow (A \vee B)$$

$$A \leftrightarrow B \equiv (A \vee B) \rightarrow (A \wedge B)$$

Notiamo che le equivalenze stabilite nel precedente esempio sono, come era da aspettarsi, tutte equazioni vere in un’algebra di Boole. Infatti, se costruiamo il quoziente dell’insieme degli enunciati rispetto alla relazione di equivalenza logica e su di esso riportiamo i connettivi come operazioni tra classi, otteniamo una vera e propria algebra di Boole, dove la classe $[A]$ precede la classe $[B]$ se $A \rightarrow B$ è sempre vero (esercizio). Di fatto,

proseguendo nella pratica di considerare gli oggetti matematici a meno di equivalenze, considereremo sempre classi di enunciati a meno dell'equivalenza logica.

Definizione 8.2. *Un enunciato si dice soddisfacibile se assume il valore di verità V per almeno un'interpretazione; in tale caso, questa interpretazione si dice modello per l'enunciato considerato.*

Definizione 8.3. *Un enunciato P che assume il valore di verità V per ogni interpretazione si dice enunciato valido o tautologia.*

Se P è un enunciato valido (tautologia), si scrive $\models P$.

Definizione 8.4. *Un enunciato si dice insoddisfacibile se non assume il valore di verità V per alcuna interpretazione, cioè*

se in ogni interpretazione assume il valore di verità F .

Un enunciato è falsificabile se assume il valore di verità F in almeno un'interpretazione.

L'enunciato dell'[Esempio 8.5](#), è una tautologia (e quindi è soddisfacibile), mentre quello dell'[Esempio 8.4](#) è soddisfacibile (ma non è una tautologia). Inoltre, dalle definizioni introdotte, segue che un enunciato P è valido (è una tautologia) se e solo se $\neg P$ è insoddisfacibile e che P è soddisfacibile se e solo se $\neg P$ è falsificabile.

8.1.5 Principii logici e ragionamento per assurdo

Naturalmente, le tautologie della logica degli enunciati sono verità nel linguaggio oggetto, ma spesso vengono assunte come

principii nel metalinguaggio, almeno quando quest’ultimo è pensato in maniera rigorosa e non colloquiale, cioè abbastanza formalizzato come il linguaggio oggetto. Ad esempio, il fatto che $A \vee \neg A$ sia una tautologia porta, a livello del metalinguaggio, al *principio del terzo escluso* (“*un’affermazione o la sua negazione deve essere per forza vera*”), mentre il fatto che $A \wedge \neg A$ sia una contraddizione porta, a livello del metalinguaggio, alla formulazione del *principio di non contraddizione* (“*un’affermazione e la sua negazione non possono essere contemporaneamente vere*”).

Utilizzeremo spesso un altro principio nelle dimostrazioni: dalla tautologia $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$, possiamo ricavare che “*ricavare B da A equivale a dire che, negare B porta necessariamente a negare*

A ”. Potremmo chiamare questo *principio di contrapposizione*, ma possiamo anche derivarne una forma più nota e più usata che è la *riduzione all’assurdo*. Se, infatti, B fosse una contraddizione (assurdo), $\neg B$ sarebbe una tautologia e, perciò, $\neg A$ sarebbe inferita da una tautologia, risultando essa stessa una tautologia (esercizio). Perciò “*provare che A è contraddittoria ($\neg A$ valida) può ridursi a provare che, supponendo A , si arriva ad un assurdo B* ”.

Esempio 8.8. *Come esempio di uso di questo principio in matematica possiamo considerare la dimostrazione del teorema:*

“*Supponendo le consuete proprietà dei sistemi numerici, si ha che $\sqrt{2}$ è un numero irrazionale*”.

Supponiamo che $\sqrt{2}$ sia un numero razionale $\frac{m}{n}$, allora $\frac{m^2}{n^2} = 2$ cioè $m^2 = 2n^2$; ma questo significa che il numero dei fattori uguali a 2 sulla sinistra dell'identità è uguale al numero di quelli a destra dell'identità. Il primo dei due tuttavia è un numero pari, il secondo dispari, quindi esiste un numero naturale pari e dispari. Quest'ultima affermazione equivale alla negazione dell'ipotesi che supponeva le consuete proprietà dei sistemi numerici.

8.2 Il metodo dei tableau proposizionali

8.2.1 La confutazione di un enunciato composto

La verifica della soddisfabilità di un enunciato attraverso la sua tavola di verità non è in generale una procedura

efficiente: essa richiede in generale 2^n calcoli, se n è il numero degli enunciati atomici che compaiono in esso. Sono stati quindi messi a punto altri metodi che, se pure non risolvono sostanzialmente il problema della complessità, possono spesso risultare più efficienti.

Il metodo dei *tableau proposizionali*¹ è originariamente un procedimento per *provare* la soddisfabilità di un enunciato. Esso, infatti, decompone l'enunciato fino alle sue componenti elementari alle quali verrà richiesto di essere vere o false per soddisfare (se possibile) l'enunciato. In questo modo il metodo fornisce anche le eventuali interpretazioni che ne saranno un modello. Il metodo viene però spesso usato per *confutare* un enunciato. Confutare un enunciato, cioè provare che esso è insoddisfacibile, significa

dimostrare che esso è falso qualsiasi siano i valori di verità degli enunciati componenti. Con il metodo dei tableau, si potrà confutare un enunciato cercando di trovare una contraddizione nella prova della sua soddisfabilità, utilizzando cioè la tecnica di riduzione all'assurdo.

Ovviamente a questo punto il metodo dei tableau potrà essere utilizzato anche per provare che un enunciato composto è una tautologia: basterà confutare la negazione dell'enunciato in esame.

Confutare $\neg A$ (cioè, provare che $\neg A$ è sempre falso) equivale a dimostrare che A è una tautologia (cioè, che A è sempre vero). Il metodo può risultare più agile, dal punto di vista esecutivo, del metodo delle tavole di verità. Osserviamo invece che non possiamo servircene per esaminare i singoli valori di verità assunti da enunciati composti (suscettibili di

assumere entrambi i valori di verità), al variare dei valori di verità assunti dagli enunciati componenti: per condurre un simile esame, non possiamo che affidarci al metodo (spesso più complicato) delle tavole di verità.

I tableau proposizionali sono grafi ad albero costituiti da una disposizione piana di nodi, contenenti uno o più enunciati; il primo nodo (detto *radice* dell'albero) contiene sempre l'enunciato in esame. A partire da questo viene costruita, attraverso regole di eliminazione dei connettivi, una tabella ramificata con nodi etichettati da enunciati sempre meno complessi, fino a giungere ai singoli enunciati componenti o i loro negati; l'idea è che gli enunciati sullo stesso nodo vanno pensati in una congiunzione metalinguistica. La costruzione ha termine quando tutti gli ultimi nodi dei

rami contengono solamente enunciati atomici o loro negazioni.

Procederemo nel modo seguente: sappiamo che la coppia di enunciati A e $\neg A$ non è simultaneamente soddisfacibile, perchè uno dei due sarà vero se e soltanto se l'altro sarà falso; la presenza nello stesso nodo di un enunciato e della sua negazione formalizza dunque un'inevitabile situazione di contraddittorietà. Ciò rende inutile procedere nell'analisi e il ramo al quale il nodo appartiene viene detto *chiuso*.

Consideriamo ora in generale i connettivi \wedge e \vee : in quale caso possiamo dire che $X \wedge Y$ è vero? Occorre che ambedue gli enunciati X e Y siano veri. In quale caso, invece, possiamo dire che $X \vee Y$ è vero? Se e solo se almeno uno tra X ed Y è vero. Dunque, nel caso di $X \wedge Y$ dobbiamo costruire un nuovo nodo nel

quale collocare entrambi gli enunciati componenti X e Y , per richiedere la verità contemporanea di entrambi; nel caso di $X \vee Y$ dobbiamo creare una biforcazione del grafo, su uno dei due nodi generati porre X e sull'altro Y , perchè si vengono a creare due situazioni, ugualmente accettabili per i nostri scopi: una che richiede soltanto la verità di X , l'altra soltanto quella di Y .

Anticipiamo che le regole che si riferiranno al connettivo \wedge determineranno l'aggiunta di un (singolo) nodo al tableau e saranno dette *α -regole*; le regole che si riferiranno a \vee determineranno la biforcazione del tableau (e dunque l'aggiunta di due nodi) e saranno dette *β -regole*.

Tali regole dovranno essere operativamente interpretate nel modo seguente: se tra gli enunciati di un nodo c'è quello presente nella prima riga della

regola, allora al ramo in questione si può aggiungere un nodo (o una coppia di nodi, nel caso delle biforazioni previste) in cui l'enunciato sia sostituito come indicato nell'ultima riga della regola. Quindi, la presenza di un enunciato nel ramo in esame (l'enunciato scritto nella prima riga) ci consente di aggiungere al ramo stesso:

- o *un unico nodo*, con uno o con due enunciati, la falsità di uno dei quali comporta la falsità dell'enunciato di partenza (regole di tipo α);
- o *due nodi* (quindi con una biforcazione), la falsità di entrambi i quali comporta la falsità dell'enunciato di partenza (regole di tipo β).

Pertanto, se all'ultimo nodo di un ramo appartengono contemporaneamente un enunciato A e la sua negazione $\neg A$,

significa che sia la falsità di A che la verità di A (quindi, la falsità di $\neg A$) comportano la falsità dell'enunciato della radice, cioè dell'enunciato da soddisfare. Si dice allora che il ramo è *chiuso*; quando *tutti* i rami sono chiusi, il tableau è chiuso e la confutazione è completata, perchè esclude ogni soddisfacibilità.

8.2.2 La costruzione di un tableau proposizionale

Formalizziamo dunque le regole per la costruzione di un tableau proposizionale, iniziando con l'evidenziare la corrispondenza dei connettivi che fanno riferimento al connettivo \wedge (ciò porterà a delle α -regole) o al connettivo \vee (ciò porterà a delle β -regole):

Regola ($\alpha 1$): riguarda $X \wedge Y$ e stabilisce che

$$\begin{array}{c} X \wedge Y \\ | \\ X, Y \end{array}$$

Regola (a2): riguarda $\neg(X \vee Y)$, cioè $\neg X \wedge \neg Y$, e stabilisce che

$$\begin{array}{c} \neg(X \vee Y) \\ | \\ \neg X, \neg Y \end{array}$$

Regola (a3): riguarda $\neg(X \rightarrow Y)$, cioè $X \wedge \neg Y$, e stabilisce che

$$\begin{array}{c} \neg(X \rightarrow Y) \\ | \\ X, \neg Y \end{array}$$

Regola (a4): riguarda $X \leftrightarrow Y$, cioè $(X \rightarrow Y) \wedge (Y \rightarrow X)$, e stabilisce che

$$\begin{array}{c}
 X \leftrightarrow Y \\
 | \\
 X \rightarrow Y, Y \rightarrow X
 \end{array}$$

Regola (α_5): riguarda $\neg\neg X$, cioè X (questa è una regola un po' diversa dalle altre perchè non riguarda un connettivo binario), e stabilisce che

$$\begin{array}{c}
 \neg\neg X \\
 | \\
 X
 \end{array}$$

Regola (β_1): riguarda $X \vee Y$ e stabilisce che

$$\begin{array}{c}
 X \vee Y \\
 / \backslash \\
 X Y
 \end{array}$$

Regola (β_2): riguarda $\neg(X \wedge Y)$, cioè $\neg X \vee \neg Y$, e stabilisce che

$$\begin{array}{c} \neg(X \wedge Y) \\ / \backslash \\ \neg X \neg Y \end{array}$$

Regola (β_3): riguarda $X \rightarrow Y$, cioè $\neg X \vee Y$, e stabilisce che

$$\begin{array}{c} X \rightarrow Y \\ / \backslash \\ \neg X Y \end{array}$$

Regola (β_4): riguarda $\neg(X \leftrightarrow Y)$, cioè $\neg(X \rightarrow Y) \vee \neg(Y \rightarrow X)$, e stabilisce che

$$\begin{array}{c} \neg(X \leftrightarrow Y) \\ / \backslash \\ \neg(X \rightarrow Y) \neg(Y \rightarrow X) \end{array}$$

È opportuno contrassegnare (ad esempio con “ \diamond ”) gli ultimi nodi dei rami chiusi.

Esempio 8.9. *Confutiamo $\neg(((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C))$. Applichiamo innanzitutto la regola (α_3) e otteniamo*

$$\begin{array}{c} \neg(((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)) \\ | \\ (A \rightarrow B) \wedge (B \rightarrow C), \neg(A \rightarrow C) \end{array}$$

Applichiamo la regola (α_1) al primo dei due enunciati ottenuti:

$$\begin{array}{c} \neg(((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)) \\ | \\ (A \rightarrow B) \wedge (B \rightarrow C), \neg(A \rightarrow C) \\ | \\ A \rightarrow B, B \rightarrow C, \neg(A \rightarrow C) \end{array}$$

e quindi la regola (α_3) all'enunciato

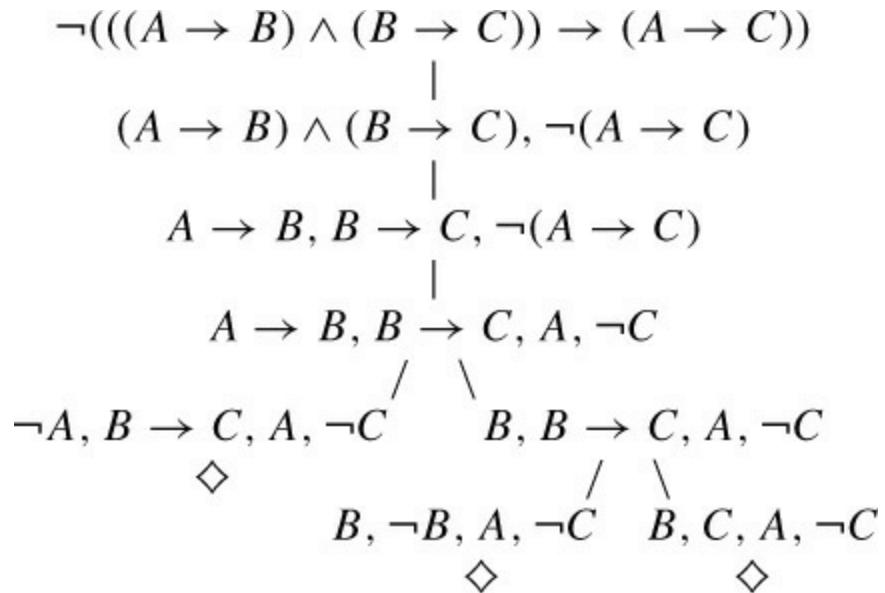
$\neg(A \rightarrow C)$:

$$\begin{array}{c} \neg(((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)) \\ | \\ (A \rightarrow B) \wedge (B \rightarrow C), \neg(A \rightarrow C) \\ | \\ A \rightarrow B, B \rightarrow C, \neg(A \rightarrow C) \\ | \\ A \rightarrow B, B \rightarrow C, A, \neg C \end{array}$$

Applichiamo la regola (β_3) al primo enunciato dell'ultimo nodo (introducendo così una biforcazione nel tableau):

$$\begin{array}{c} \neg(((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)) \\ | \\ (A \rightarrow B) \wedge (B \rightarrow C), \neg(A \rightarrow C) \\ | \\ A \rightarrow B, B \rightarrow C, \neg(A \rightarrow C) \\ | \\ A \rightarrow B, B \rightarrow C, A, \neg C \\ / \quad \backslash \\ \neg A, B \rightarrow C, A, \neg C \quad B, B \rightarrow C, A, \neg C \\ \diamond \end{array}$$

Il ramo di sinistra è chiuso, in quanto nell'ultimo nodo troviamo sia A che $\neg A$; possiamo abbandonarne l'esame e proseguire la formazione del tableau con il solo ramo a destra. Applichiamo la regola $(\beta 3)$ a $B \rightarrow C$ (e ciò provoca un'ulteriore biforcazione):



I due rami formati sono entrambi chiusi: quello a sinistra per la presenza contemporanea di $\neg B$ e B nell'ultimo nodo; quello a destra per la presenza contemporanea di C e $\neg C$ nell'ultimo

nodo.

Tutti i rami del tableau risultano dunque chiusi: l'enunciato di partenza è confutato e ciò significa che la sua negazione, cioè $((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)$, è una tautologia (talvolta detta legge del sillogismo ipotetico).

Esempio 8.10. *Esaminiamo la formula $P \wedge (\neg Q \vee \neg P)$. Proviamo a costruire il tableau della formula data (non negata):*

$$\begin{array}{c} P \wedge (\neg Q \vee \neg P) \\ | \\ P, \neg Q \vee \neg P \\ / \quad \backslash \\ P, \neg Q \quad P, \neg P \\ \diamond \end{array}$$

Non abbiamo ottenuto un tableau chiuso: il ramo a destra è infatti chiuso, ma quello a sinistra no. Ciò significa che la formula data è soddisfacibile; inoltre, dal ramo aperto ci viene anche fornita l'interpretazione che la rende soddisfatta:

$$v(P) = V \text{ e } v(Q) = F$$

Esempio 8.11. Proviamo allora a costruire il tableau della negazione della formula esaminata nell'esempio precedente:

$$\begin{array}{c} \neg(P \wedge (\neg Q \vee \neg P)) \\ \diagup \quad \diagdown \\ \neg P \quad \neg(\neg Q \vee \neg P) \end{array}$$

Possiamo fermarci qui, perché il ramo di sinistra non si chiuderà mai. Infatti, per $v(P) = F$, la formula negata è soddisfacibile.

Attenzione! Non è vero in generale che, se tutti i rami restano aperti, la formula dalla quale si è partiti risulta una tautologia: significa soltanto che è soddisfatta per tutti i valori specificati dai rami aperti. Questi, però, in generale non esauriscono tutti i casi possibili come accadeva nella verifica con le tavole di verità.

8.2.3 Correttezza e completezza

Abbiamo più volte affermato che il metodo dei tableau proposizionali consente di stabilire se una formula è valida ma, per il momento, si è trattato soltanto di una giustificazione intuitiva. Ci accingiamo ora a provare formalmente che questo è il caso.

Si può provare innanzitutto che la costruzione di un tableau proposizionale di un enunciato, condotta secondo il procedimento precedentemente descritto, termina dopo un numero *finito* di passi e che su ogni foglia abbiamo soltanto enunciati atomici o loro negazioni (detti anche *letterali*). Ciò è facilmente dimostrabile per induzione strutturale, perché supponendo di non utilizzare il connettivo \leftrightarrow (d'altronde facilmente simulabile con una combinazione di altri), l'applicazione delle regole porta ad ogni

passo ad una sostanziale diminuzione dei connettivi presenti nel nodo e questi sono in numero finito. Questa dimostrazione, che si lascia al lettore come utile esercizio, è un esempio di applicazione dell'induzione strutturale: si associa ad ogni enunciato il numero dei suoi connettivi binari moltiplicato per 2 più il numero delle negazioni che vi compaiono. Si vedrà subito che, in questo modo, ad ogni passo della costruzione del tableau, la somma degli indici associati all'insieme degli enunciati che compaiono nel nodo, o nei nodi, che si vengono a creare decresce e, dal momento che non può scendere al di sotto dello 0, il procedimento deve terminare.

Si prova poi che, detto T un tableau completo per P , P è insoddisfacibile se e soltanto se il tableau T è chiuso.
L'implicazione “se T è chiuso allora P è

insoddisfacibile" corrisponde alla *correttezza* del metodo, perchè significa che ciò che esso prova falso (vero) è falso (vero) anche rispetto alle tavole di verità. Di conseguenza, un metodo corretto è quello che non commette errori di valutazione. Questo potrebbe farlo anche un metodo banale che non provasse nulla. Noi, però, dimostreremo anche l'implicazione inversa: "se P è *insoddisfacibile* allora T è chiuso" corrisponde alla *completezza* del metodo; cioè con questo metodo riusciamo a confutare (verificare) tutte le proposizioni false (vere) rispetto alle tavole di verità. Un sistema è completo se è abbastanza potente da riuscire a provare tutto quello che c'è da provare. Anche qui si potrebbe pensare ad un metodo banale che provasse tutto. Questo sarebbe completo, ma non corretto, perchè proverebbe cose false, risultando *incoerente*.

Teorema 8.1 (Correttezza e completezza del metodo dei tableau). *La formula P è valida (è una tautologia) se e soltanto se il tableau per $\neg P$ è chiuso.*

Dimostrazione.

(Correttezza) Dimostreremo un fatto più generale: se un sottoalbero radicato nel nodo n del tableau T è chiuso, allora l'insieme di formule $U(n)$ in n è insoddisfacibile. La dimostrazione è per induzione sull'altezza h del nodo n nel tableau considerato.

Base: Se $h = 0$ e T è chiuso, allora il nodo contiene due enunciati che sono uno la negazione dell'altro e pertanto $U(n)$ è insoddisfacibile.

Passo induttivo: Se $h > 0$, allora è stata utilizzata qualche regola di tipo α per un

connettivo riconducibile a \wedge o di tipo β per un connettivo riconducibile a \vee .

Distinguiamo i due casi:

- Nel caso di una regola α , si ha che $U(n) = \{P_1 \wedge P_2\} \cup U_o$ e $U(n)' = \{P_1, P_2\} \cup U_o$, dove n' è il figlio di n' in T e U_o può essere vuoto. Chiaramente, l'altezza di n' è $h - 1$; dunque, per l'ipotesi induttiva, $U(n)'$ è insoddisfacibile. Quindi, se indichiamo con v una qualsiasi interpretazione, deve essere $v(P') = F$ per qualche $P' \in U(n)'$. Abbiamo tre possibilità:
 - per qualche $P_o \in U_o$ è $v(P_o) = F$; ma è $P_o \in U_o$;
 - $v(P_1) = F$; allora $v(P_1 \wedge P_2) = F$;
 - $v(P_2) = F$; ancora $v(P_1 \wedge P_2) = F$.

In tutti e tre i casi abbiamo trovato un $P \in U(n)$ tale che $v(P) = F$; quindi, $U(n)$ è insoddisfacibile.

- Nel caso di una regola β , si ha che $U(n) = \{P_1 \vee P_2\} \cup U_0$, $U(n') = \{P_1\} \cup U_0$ e $U(n'') = \{P_2\} \cup U_0$, dove n' ed n'' sono i due figli di n in T ; per l'ipotesi induttiva, sia $U(n')$ che $U(n'')$ sono insoddisfacibili. Quindi, indicando ancora con v una qualsiasi interpretazione, abbiamo due possibilità:

1. per qualche $P_0 \in U_0$ è $v(P_0) = F$; ma è $P_0 \in U_0 \subseteq U(n)$;
2. se invece è $v(P_0) = V$ per ogni $P_0 \in U_0$, affinchè sia $U(n')$ che $U(n'')$ siano insoddisfacibili, deve essere $v(P_1) = v(P_2) = F$, quindi $v(P_1 \vee P_2) = F$.

Da ciò possiamo nuovamente concludere che $U(n)$ è insoddisfacibile.

(Completezza) Per dimostrare che se P

è insoddisfacibile allora il tableau per P è chiuso, proveremo che se in tale tableau ci fosse un ramo aperto allora P sarebbe soddisfacibile. Procediamo ancora per induzione sull'altezza h del nodo n nel tableau considerato.

Base: Se $h = 0$ e il tableau è aperto, allora il nodo non contiene due letterali che sono uno la negazione dell'altro e P è soddisfacibile.

Passo induttivo: Se $h > 0$, allora è stata utilizzata qualche regola di tipo α o di tipo β . Nel caso delle regole α , ragionando come precedentemente fatto nel caso della correttezza, esiste un'interpretazione v tale che $v(P') = V$ per ogni $P \in U(n')$ e $U(n)$ è soddisfacibile. Nel caso delle regole β , esiste un'interpretazione v tale che per ogni $P_0 \in U_0$ è $v(P_0) = V$ e che almeno

uno tra $v(P_1)$ e $v(P_2)$ sia V ; da ciò soddisfacibile. segue che $v(P_1 \vee P_2) = V$. Dunque, $U(n)$ è

□

8.3 Il sistema deduttivo di Gentzen

8.3.1 Il sistema G

Finora abbiamo considerato due metodi, le tavole di verità ed i tableau, che permettono di *verificare* la soddisfabilità/validità di un enunciato calcolandone il valore di verità per tutte o alcune interpretazioni. Questi metodi fanno pesantemente appello dunque alla semantica.

Questa procedura, anche se nei casi semplici può far comodo, deve essere

superata se vogliamo imparare metodi formali per fare *dimostrazioni* come quelle cui siamo abituati, ad esempio, in geometria. Per fare questo abbiamo bisogno di tecniche totalmente sintattiche che, partendo da enunciati o famiglie di enunciati che assumiamo valere e applicando opportune regole, ci portino a qualcosa che dovrà ancora valere. La garanzia di questo fatto ci verrà fornita dal solito (meta)teorema di correttezza. Si noti che stiamo sempre di più prescindendo dal riferimento all'interpretazione delle formule, cioè al valore di verità che esse assumono; ci stiamo muovendo, cioè, a livello sempre più *sintattico*, cioè guardando alla struttura delle formule, piuttosto che alla loro *semantica*, che, in questo caso, è data dai valori di verità. Questo fatto servì a Euclide, come più tardi ad Hilbert, per acquisire un maggior rigore.

Un sistema deduttivo in generale serve a provare una formula mediante un teorema. Un *teorema* è una successione *finita* ed ordinata di insiemi di formule del linguaggio costruita in modo tale che ognuno di essi sia un *assioma* o sia ottenuto dai precedenti attraverso le *regole di inferenza* del sistema.

Presenteremo prima di tutti quello che chiameremo sistema di Gentzen² (o sistema *G*), anche se non coincide esattamente con nessuno dei sistemi che Gentzen propose, perché utilizza soltanto regole di inserzione di connettivi.

Seguiamo in questa scelta l'approccio di [6], perchè in questo modo viene messo in risalto il suo legame con il metodo dei tableau, che risulta essere una vera e propria dualità. Sarà allora quasi immediato dimostrare il teorema di correttezza e completezza di questo sistema rispetto ai tableau.

Nel caso del sistema G , un assioma è un insieme finito di formule U che contiene un enunciato e la sua negazione (P e $\neg P$). Gli assiomi sono quegli insiemi di formule che dovranno essere considerati necessariamente veri; nel nostro caso la scelta è dovuta al fatto che mettere insieme degli enunciati in un insieme significa qui considerarli intuitivamente in una disgiunzione metalinguistica; perciò, la presenza di una coppia ($P, \neg P$) ci garantisce la soddisfabilità in ogni interpretazione. Siamo esattamente nella situazione duale a quella del metodo dei tableau, dove mettere insieme due enunciati significava considerarli in una congiunzione metalinguistica: lì, la presenza di una coppia ($P, \neg P$) rappresentava l'insoddisfabilità.

Di conseguenza il sistema G si presenta in maniera del tutto speculare rispetto a

quello dei tableau. Nel metodo dei tableau proposizionali abbiamo considerato delle regole α (quelle che non introducono biforazioni nel tableau considerato) e delle regole β (che introducono una biforazione). Faremo un’analoga distinzione anche nel caso delle regole di inferenza per il sistema G. Stavolta, però, le regole serviranno ad introdurre i connettivi, invece di eliminarli e le regole α saranno legate ai connettivi del tipo \vee , mentre le regole β saranno legate ai connettivi del tipo \wedge .

Considereremo regole basate sulle tabelle seguenti:

| α | α_1 | α_2 | β | β_1 | β_2 |
|-----------------------------|-------------------------|-------------------------|-------------------------|-------------------|-------------------|
| $P \vee Q$ | P | Q | $P \wedge Q$ | P | Q |
| $\neg(P \wedge Q)$ | $\neg P$ | $\neg Q$ | $\neg(P \vee Q)$ | $\neg P$ | $\neg Q$ |
| $P \rightarrow Q$ | $\neg P$ | Q | $\neg(P \rightarrow Q)$ | P | $\neg Q$ |
| $\neg(P \leftrightarrow Q)$ | $\neg(P \rightarrow Q)$ | $\neg(Q \rightarrow P)$ | $P \leftrightarrow Q$ | $P \rightarrow Q$ | $Q \rightarrow P$ |
| P | $\neg\neg P$ | | | | |

Le regole di inferenza sono dei due tipi seguenti: la regola di inferenza relativa alla tabella di tipo α è

$$\frac{U \cup \{\alpha_1, \alpha_2\}}{U \cup \{\alpha\}}$$

mentre la regola di inferenza relativa alla tabella di tipo β è

$$\frac{U_1 \cup \{\beta_1\} \quad U_2 \cup \{\beta_2\}}{U_1 \cup U_2 \cup \{\beta\}}$$

Una dimostrazione nel sistema G è una successione di insiemi di formule tale che ciascun elemento della successione o è un assioma o può essere inferito da elementi precedenti attraverso le regole. L'ultimo elemento P è detto *dimostrabile*, scritto $\vdash_G P$ o, più semplicemente, scritto $\rightarrow P$.

Esempio 8.12. *Dimostriamo nel sistema G il teorema $\rightarrow(A \vee B) \rightarrow (B \vee A)$. La dimostrazione è la seguente:*

1. $A, B, \neg A$ (assioma);
2. $B, \neg B, A$ (assioma);
3. $\neg(A \vee B), B, A$ (regola β con premessa i punti 1 e 2);
4. $\neg(A \vee B), (B \vee A)$ (regola α con premessa il punto 3);
5. $(A \vee B) \rightarrow (B \vee A)$ (regola α con premessa il punto 4).

8.3.2 Deduzione in G e tableau

Una deduzione nel sistema di Gentzen può essere posta in forma di albero: anticipiamo che tale possibilità si rivelerà interessante per il collegamento che essa consentirà con il metodo dei tableau.

Esempio 8.13. Scriviamo la dimostrazione dell'esempio precedente in maniera più schematica:

$$\begin{array}{c}
 \dfrac{\neg A, B, A \quad \neg B, B, A}{\neg(A \vee B), B, A} \\
 \dfrac{}{\neg(A \vee B), (B \vee A)} \\
 \dfrac{}{(A \vee B) \rightarrow (B \vee A)}
 \end{array}$$

o, equivalentemente, in forma di albero:

$$\begin{array}{c}
 \begin{array}{ccc}
 \neg A, B, A & \neg B, B, A & \\
 \backslash & / & \\
 \neg(A \vee B), B, A & & \\
 | & & \\
 \neg(A \vee B), (B \vee A) & & \\
 | & & \\
 (A \vee B) \rightarrow (B \vee A) & &
 \end{array}
 \end{array}$$

In quest'ultima scrittura risulta evidente il ruolo della regola di inferenza di tipo β (corrispondente alla biforcazione) e delle due regole di tipo α , applicate successivamente.

Al lettore non sfuggirà l'analogia di tale disposizione di formule con il tableau (graficamente “capovolto”!) che potrebbe essere costruito per la negazione della formula $(A \vee B) \rightarrow (B \vee A)$. Nell'esempio seguente costruiremo tale tableau.

Esempio 8.14. Il tableau della negazione di $(A \vee B) \rightarrow (B \vee A)$ è:

$$\begin{array}{c}
 \neg((A \vee B) \rightarrow (B \vee A)) \\
 | \\
 A \vee B, \neg(B \vee A) \\
 | \\
 A \vee B, \neg B, \neg A \\
 / \quad \backslash \\
 A, \neg B, \neg A \quad B, \neg B, \neg A \\
 \diamond \quad \diamond
 \end{array}$$

Consideriamo la deduzione in G della formula $(A \vee B) \rightarrow (B \vee A)$ che abbiamo ricavato nell'esempio precedente. In essa compaiono formule simili a quelle che compaiono in questo tableau: in particolare, le formule del tableau sono esattamente le negazioni delle formule che compaiono nell'albero che rappresenta la deduzione in Gentzen. In particolare, gli assiomi dai quali la precedente deduzione nel sistema di Gentzen trae origine possono essere ritrovati, con le negazioni scambiate, nei nodi finali del tableau che consentono la

chiusura dei rami:

$A, \neg B, \neg A \wedge B, \neg B, \neg A$ come chiusure dei rami del tableau;

$\neg A, B, A \wedge \neg B, B, A$ come assiomi di Gentzen.

Esempio 8.15. Dimostriamo nel sistema G il teorema $\neg((A \vee B) \wedge (\neg A \wedge \neg B))$. La deduzione è la seguente:

$$\begin{array}{c} \frac{\neg A, A, B \quad \neg B, A, B}{\neg(A \vee B), A, B} \\ \frac{}{\neg(A \vee B), \neg\neg A, \neg\neg B} \\ \frac{}{\neg(A \vee B), \neg(\neg A \vee \neg B)} \\ \frac{}{\neg((A \vee B) \wedge (\neg A \wedge \neg B))} \end{array}$$

Costruiamo ora il tableau della negazione di $\neg((A \vee B) \wedge (\neg A \wedge \neg B))$, ovvero il tableau di $(A \vee B) \wedge (\neg A \wedge \neg B)$:

$$\begin{array}{c}
 (A \vee B) \wedge (\neg A \wedge \neg B) \\
 | \\
 (A \vee B), (\neg A \wedge \neg B) \\
 | \\
 (A \vee B), \neg A, \neg B \\
 / \quad \backslash \\
 A, \neg A, \neg B \quad B, \neg A, \neg B \\
 \diamond \quad \diamond
 \end{array}$$

È da notare che, mentre il metodo di verifica dei tableau è totalmente meccanico, cioè, partendo da un enunciato, si procede all'eliminazione in esso dei connettivi, la deduzione in G richiede un *procedimento euristico*: come si fa a sapere quali sono gli assiomi da mettere all'inizio della dimostrazione? Normalmente si procede a ritroso, partendo dall'enunciato da dimostrare e risalendo agli assiomi, chiedendosi ogni volta quali regole si sarebbero potute applicare.

Esempio 8.16. Riprendiamo l'esempio precedente. Dovendo dimostare come teorema l'enunciato $\neg((A \vee B) \wedge (\neg A \wedge \neg B))$

$\neg B$), possiamo pensare che l'ultimo passo della dimostrazione sarà stata l'applicazione di un'a regola alla coppia di enunciati $\neg(A \vee B)$ e $\neg(\neg A \wedge \neg B)$. Il secondo di questi potrebbe essere stato ottenuto mediante l'applicazione di un'a regola alla coppia di enunciati $\neg\neg A$ e $\neg\neg B$; ricostruiamo in questo modo il passaggio che ci porta a $\neg(A \vee B)$, $\neg\neg A$, $\neg\neg B$. Quest'ultimo è a sua volta facilmente ottenibile da $\neg(A \vee B)$, A , B . A questo punto, l'ormai prevedibile applicazione di una regola β ci dà gli assiomi $\neg A$, A , B e $\neg B$, A , B

La stessa cosa accadrà nel sistema di Hilbert che illustreremo nel prossimo paragrafo.

Siamo ora in grado di provare il teorema di completezza e di correttezza del sistema di Gentzen rispetto al metodo dei tableau; per transitività, esso varrà

anche rispetto alle tavole di verità.

Teorema 8.2 (Correttezza e completezza del sistema G). *Una formula è da se e soltanto se è dimostrabile nel sistema G .*

In questo caso la dimostrazione, che viene lasciata come esercizio, può procedere rilevando che P è dimostrabile in Gentzen se e soltanto se $\neg P$ ha un tableau chiuso. La prova è una semplicissima induzione sull'altezza dell'albero di dimostrazione che viene trasformato in un tableau chiuso invertendo i passaggi e negando le formule e viceversa.

8.4 Il sistema deduttivo di Hilbert

8.4.1 Assiomi e Modus Ponens

Consideriamo ora un altro sistema deduttivo, che ha preceduto storicamente quello di Gentzen e che è più vicino al nostro modo di fare dimostrazioni in matematica. Esso costituitì infatti il tentativo di dare una formalizzazione così rigorosa al modo di ricavare teoremi in matematica da porre quest'ultima al riparo da tutte le possibili antinomie. Tale scopo non è stato di fatto raggiunto, perché un risultato di Gödel prova che è impossibile garantire alla matematica la non contraddittorietà all'interno della stessa matematica. Il sistema H o di Hilbert³ [24] è tuttavia usato tuttora come principale sistema deduttivo.

In questo sistema, almeno nella forma che qui proponiamo, abbiamo sempre a che fare non con singoli enunciati, ma con *schemi* di enunciati, cioè enunciati dati *a meno di sostituzioni uniformi degli atomi*

con enunciati. Questo vuol dire che si può sempre sostituire ogni lettera con un nuovo enunciato, purchè lo si faccia in modo uniforme: a lettera uguale deve corrispondere uguale sostituzione. Da quanto finora detto, è chiaro che le proprietà di validità di un enunciato si conservano per sostituzione uniforme, cioè una tautologia resta tale, ma qualcosa che non lo è può diventarlo.

Una dimostrazione in H sarà una successione finita di enunciati tale che ognuno di essi o è l'istanza di un assioma o è ottenuto da due enunciati precedenti attraverso il Modus Ponens.

Esempio 8.17. *La proposizione $A \vee \neg A$ è una tautologia e resta tale se sostituiamo ovunque A con un qualsiasi enunciato, ad esempio $B \rightarrow C$: è infatti facile verificare che $(B \rightarrow C) \vee \neg(B \rightarrow C)$ è una tautologia. D'altra parte, l'enunciato*

$A \vee \neg B$ non è una tautologia ma lo diventa se sostituiamo sia A che B con l'enunciato $B \rightarrow C$.

Dal momento che nel sistema di Hilbert ci occuperemo di teoremi (che poi, grazie alla correttezza ed alla completezza, coincideranno con le tautologie), possiamo lavorare a meno di sostituzioni uniformi. Una qualunque sostituzione uniforme di proposizioni a lettere nelle seguenti formule è un assioma nel sistema di Hilbert (sistema H), dove $\vdash H$ (o, più semplicemente, \vdash) esprime, al solito, la dimostrabilità:

| | |
|-------------|--|
| (Assioma 1) | $\vdash A \rightarrow (B \rightarrow A)$ |
| (Assioma 2) | $\vdash (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$ |
| (Assioma 3) | $\vdash (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$ |

Osserviamo innanzitutto che ci sono tre schemi ma infiniti assiomi, perchè A e B possono essere a loro volta sostituite con

qualsiasi formula. La regola di inferenza nel sistema di Hilbert è detta *Modus Ponens* (MP) e si formalizza nel modo seguente:

$$\frac{\vdash A \quad \vdash A \rightarrow B}{\vdash B}$$

Esempio 8.18. Dimostriamo nel sistema H il teorema $\vdash A \rightarrow A$. La dimostrazione è la seguente, tenendo conto del fatto che possiamo utilizzare qualunque istanza dello stesso schema di assioma:

1. Partiamo dall'Assioma 2, dove abbiamo sostituito $A \rightarrow A$ a B e A a C :

$$\vdash (A \rightarrow ((A \rightarrow A) \rightarrow A)) \rightarrow ((A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A))$$

2. Prendiamo l'istanza dell'Assioma 1 ottenuta sostituendo $A \rightarrow A$ a B :

$$\vdash A \rightarrow ((A \rightarrow A) \rightarrow A)$$

3. Per modus ponens tra 1 e 2 abbiamo

$$\vdash ((A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A))$$

4. Prendiamo l'istanza dell'Axioma 1
ottenuta sostituendo A a B :

$$\vdash A \rightarrow (A \rightarrow A)$$

5. Per modus ponens tra 3 e 4 abbiamo

$$\vdash A \rightarrow A$$

Ci si può legittimamente domandare come si sia arrivati ad immaginare un tale tipo di dimostrazione. Chiaramente non esiste un modo meccanico di generarla. Potremmo pensare di aver

ragionato (euristicamente) nel seguente modo: dal momento che la nostra unica regola è il Modus Ponens, l'unica speranza di dimostrare $A \rightarrow A$ sta nella possibilità di dimostrare una formula $D \rightarrow (A \rightarrow A)$, dove D sia a sua volta dimostrabile. Per prima cosa ci viene in mente di usare l'assioma 1, dove abbiamo istanziato B con A ; ma a questo punto D sarebbe A e, quindi, certamente non dimostrabile, perché A è una formula qualunque. Proviamo allora ad usare l'assioma 2 con C istanziato con A . Ci rendiamo conto allora che la premessa della formula dell'assioma 2 diventa, in questo caso, un'istanza dell'assioma 1, perciò dimostrabile e, attraverso il Modus Ponens, rendendo dimostrabile anche la conseguenza, cioè:

$$(A \rightarrow B) \rightarrow (A \rightarrow A)$$

Come dunque istanziare B in modo da poter ripetere il trucco? L'idea è fare in modo che $A \rightarrow B$ diventi istanza di un assioma; questo è possibile sostituendo B con $A \rightarrow B$ o, ancora più semplicemente, con $A \rightarrow A$, come abbiamo fatto nella nostra dimostrazione sopra riportata, ottenendo un'altra istanza dell'assioma 1.

Da questo esempio introduttivo si può notare che una formula estremamente semplice (la tesi era $\vdash A \rightarrow A$) richiede già una dimostrazione tecnicamente piuttosto complicata e difficile da immaginare. Ciò suggerisce l'opportunità di introdurre alcune *regole derivate* per il sistema H , che rendano lo stesso più maneggevole che con il solo MP. Queste regole saranno dimostrate a livello metalinguistico, cioè si proverà che una dimostrazione che le usa può essere trasformata in una dimostrazione che non le usa. Quindi le

nuove regole non aggiungeranno nulla alla potenza dimostrativa del sistema H . Questo fatto è molto importante perché aumentare la capacità dimostrativa di un sistema può renderlo incoerente, cioè non corretto.

8.4.2 Regole derivate del sistema di Hilbert

Come sopra anticipato, il sistema di Hilbert nella forma originale (avente il *Modus Ponens* come unica regola di inferenza) appare di applicazione assai ostica. Per agevolarne l'uso vengono pertanto introdotte, nel sistema H , varie *regole derivate*, la più importante delle quali è la *regola di deduzione*.

Per utilizzare questa regola avremo bisogno di generalizzare la nozione di dimostrazione. Abbiamo detto che in una dimostrazione possiamo introdurre ad

ogni passo un assioma o una formula ottenuta dalle precedenti mediante una regola del sistema; in realtà, siamo abituati dalla pratica matematica a lavorare anche con *assunzioni*, cioè formule che assumiamo vere per quella particolare dimostrazione e che, quindi, dobbiamo sempre dichiarare nella prova come un debito contratto. Ad esempio, se dimostriamo una proprietà per i triangoli isosceli, questa vale sotto l'assunzione che due lati siano uguali e non varrà in generale per tutti i triangoli. Con la scrittura

$$U \vdash A$$

indicheremo che le formule presenti in U sono assunzioni nella dimostrazione di A (se U è vuoto, vuol dire che nel dimostrare A abbiamo usato soltanto assiomi).

Ora, la regola di deduzione è espressa nel modo seguente:

$$\frac{U \cup \{A\} \vdash B}{U \vdash A \rightarrow B}$$

Essa afferma che, se abbiamo dimostrato B usando le assunzioni contenute in $U \cup \{A\}$, possiamo considerare dimostrato $A \rightarrow B$ sotto le sole assunzioni contenute in U . Nel nostro esempio, se dagli assiomi della geometria euclidea e dall'assunzione che due lati sono uguali abbiamo dimostrato la proprietà P , allora dai soli assiomi della geometria euclidea sappiamo di poter dimostrare che “se due lati sono uguali allora P ”. In questo modo, dovendo dimostrare un'implicazione, possiamo assumere la premessa e dimostrare la conseguenza: come se, usando la nostra metafora, contraessimo un prestito che alla fine restituiamo.

Spesso, ciò renderà le dimostrazioni

tecnicamente molto più facili. Infatti, dovendo in ogni dimostrazione procedere euristicamente a ritroso, partendo dall'enunciato da dimostrare e risalendo agli assiomi, chiedendosi ogni volta quali regole si sarebbero potute applicare, suggerirà il modo nel quale procedere, perchè ci metterà a disposizione delle ipotesi aggiuntive che potremo utilizzare come fossero assiomi, come si può vedere negli esempi. Infatti, sempre supponendo per semplicità U vuoto, se dall'assumere A si dimostra B (in simboli, $\{A\} \rightarrow B$), allora posso dimostrare che se A allora B (in simboli, $A \rightarrow B$); e viceversa, usando il *Modus Ponens*. Purtroppo, questa perfetta corrispondenza si perderà non appena i calcoli logici saranno più complessi.

Proposizione 8.1. *La regola di deduzione è una regola derivata corretta.*

Dimostrazione. Si procede per induzione sulla lunghezza n della dimostrazione $U \cup \{A\} \vdash B$.

Se $n = 1$, B si dimostra in un passo; dunque, B può essere A , un elemento di U oppure un assioma. Se B è A , allora $\vdash A \rightarrow B$ (si veda l'[Esempio 8.18](#)) e, dunque, $U \vdash A \rightarrow B$. Altrimenti, una dimostrazione (in cui non si usa la regola derivata) di $U \vdash A \rightarrow B$ è data da:

- | | |
|---|-------------|
| 1. $U \vdash B$ | (Ipotesi) |
| 2. $U \vdash B \rightarrow (A \rightarrow B)$ | (Assioma 1) |
| 3. $U \vdash A \rightarrow B$ | (MP 1,2) |

Se è $n > 1$, l'ultimo passo nella dimostrazione di $U \cup \{A\} \rightarrow B$ è un'inferenza del tipo precedente, oppure un'inferenza di B che usa il *Modus Ponens*. Nel primo caso il risultato si ottiene dalla dimostrazione per $n = 1$. Se è stato usato il *Modus Ponens*, allora esiste una formula C tale che la i -esima formula

nella dimostrazione è $U \cup \{A\} \vdash C$ e la j -esima formula è $U \cup \{A\} \vdash C \rightarrow B$, essendo $i < n$ e $j < n$. Mediante l'ipotesi induttiva, si ottiene una dimostrazione di $U \vdash A \rightarrow C$ e di $U \vdash A \rightarrow (C \rightarrow B)$. Una dimostrazione di $U \vdash A \rightarrow B$ è allora:

$$\begin{aligned}
 i'. U \vdash A \rightarrow C \\
 j'. U \vdash A \rightarrow (C \rightarrow B) \\
 j'+1. U \vdash (A \rightarrow (C \rightarrow B)) \rightarrow ((A \rightarrow C) \rightarrow (A \rightarrow B)) & \text{(Assioma 2)} \\
 j'+2. U \vdash (A \rightarrow C) \rightarrow (A \rightarrow B) & \text{(MP } j', j'+1) \\
 j'+3. U \vdash A \rightarrow B & \text{(MP } i', j'+2) \quad \square
 \end{aligned}$$

Possiamo allora concludere che ogni dimostrazione che usa la regola di deduzione può essere trasformata in una dimostrazione che non la usa seguendo il procedimento illustrato nella dimostrazione della precedente proposizione.

Usando il *Modus Ponens*, possiamo ancora provare facilmente da alcuni teoremi del sistema di Hilbert le seguenti regole derivate. Queste si presentano

ormai come principi logici, cioè regole che riguardano le dimostrazioni. Il procedimento sarà quello di dimostrare un teorema e ricavarne la corrispondente regola.

Proposizione 8.2 (Regola di contrapposizione).

$$\frac{U \vdash \neg B \rightarrow \neg A}{U \vdash A \rightarrow B}$$

Dimostrazione. La regola di contrapposizione può essere dedotta immediatamente dall'assioma 3:

- | | |
|---|---------------|
| 1. $U \vdash \neg B \rightarrow \neg A$ | Ipotesi |
| 2. $U \vdash (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$ | Assioma 3 |
| 3. $U \vdash A \rightarrow B$ | <i>MP</i> 1,2 |

□

Proposizione 8.3 (Regola di transitività).

$$\frac{U \vdash A \rightarrow (B \rightarrow (C \rightarrow D))}{U \vdash A \rightarrow (C \rightarrow (B \rightarrow D))}$$

Dimostrazione. La regola di transitività dell’implicazione può essere facilmente inferita una volta dimostrata la seguente tautologia:

$$\vdash (A \rightarrow B) \rightarrow [(B \rightarrow C) \rightarrow (A \rightarrow C)]$$

La sua prova in Hilbert è la seguente:

- | | | |
|----|--|-------------|
| 1. | $\{A \rightarrow B, B \rightarrow C, A\} \vdash A$ | Ipotesi |
| 2. | $\{A \rightarrow B, B \rightarrow C, A\} \vdash A \rightarrow B$ | Ipotesi |
| 3. | $\{A \rightarrow B, B \rightarrow C, A\} \vdash B$ | MP 1,2 |
| 4. | $\{A \rightarrow B, B \rightarrow C, A\} \vdash B \rightarrow C$ | Ipotesi |
| 5. | $\{A \rightarrow B, B \rightarrow C, A\} \vdash C$ | MP 3,4 |
| 6. | $\{A \rightarrow B, B \rightarrow C\} \vdash A \rightarrow C$ | Deduzione 5 |
| 7. | $\{A \rightarrow B\} \vdash [(B \rightarrow C) \rightarrow (A \rightarrow C)]$ | Deduzione 6 |
| 8. | $\vdash (A \rightarrow B) \rightarrow [(B \rightarrow C) \rightarrow (A \rightarrow C)]$ | Deduzione 7 |

□

Proposizione 8.4 (Regola di scambio delle premesse).

$$\frac{U \vdash A \rightarrow (B \rightarrow C)}{U \vdash B \rightarrow (A \rightarrow C)}$$

Dimostrazione. La dimostrazione segue facilmente, una volta dimostrata la seguente tautologia:

$$\vdash [A \rightarrow (B \rightarrow C)] \rightarrow [B \rightarrow (A \rightarrow C)]$$

la cui prova in Hilbert è la seguente:

- | | | |
|----|--|-------------|
| 1. | $\{A \rightarrow (B \rightarrow C), B, A\} \vdash A$ | Ipotesi |
| 2. | $\{A \rightarrow (B \rightarrow C), B, A\} \vdash A \rightarrow (B \rightarrow C)$ | Ipotesi |
| 3. | $\{A \rightarrow (B \rightarrow C), B, A\} \vdash B \rightarrow C$ | MP 1,2 |
| 4. | $\{A \rightarrow (B \rightarrow C), B, A\} \vdash B$ | Ipotesi |
| 5. | $\{A \rightarrow (B \rightarrow C), B, A\} \vdash C$ | MP 3,4 |
| 6. | $\{A \rightarrow (B \rightarrow C), B\} \vdash A \rightarrow C$ | Deduzione 5 |
| 7. | $\{A \rightarrow (B \rightarrow C)\} \vdash B \rightarrow (A \rightarrow C)$ | Deduzione 6 |
| 8. | $\vdash [A \rightarrow (B \rightarrow C)] \rightarrow [B \rightarrow (A \rightarrow C)]$ | Deduzione 7 |

□

Proposizione 8.5 (Regola dello Pseudo-Scoto).

$$\frac{U \vdash A \quad U \vdash \neg A}{U \vdash B}$$

Dimostrazione. La dimostrazione segue facilmente, una volta dimostrata la seguente tautologia:

$$\vdash \neg A \rightarrow (A \rightarrow B)$$

la cui prova in Hilbert è la seguente:

- | | | |
|----|--|-------------|
| 1. | $\{\neg A, A\} \vdash \neg A \rightarrow (\neg B \rightarrow \neg A)$ | Assioma 1 |
| 2. | $\{\neg A, A\} \vdash \neg A$ | Ipotesi |
| 3. | $\{\neg A, A\} \vdash \neg B \rightarrow \neg A$ | MP 1,2 |
| 4. | $\{\neg A, A\} \vdash (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$ | Assioma 3 |
| 5. | $\{\neg A, A\} \vdash A \rightarrow B$ | MP 3,4 |
| 6. | $\{\neg A, A\} \vdash A$ | Ipotesi |
| 7. | $\{\neg A, A\} \vdash B$ | MP 5,6 |
| 8. | $\{\neg A\} \vdash A \rightarrow B$ | Deduzione 8 |
| 9. | $\vdash \neg A \rightarrow (A \rightarrow B)$ | Deduzione 9 |

□

Proposizione 8.6 (Regola della doppia negazione).

$$\frac{U \vdash \neg\neg A}{U \vdash A}$$

Dimostrazione.

- | | | |
|----|--|--------------------|
| 1. | $\{\neg\neg A\} \vdash \neg\neg A \rightarrow (\neg\neg\neg A \rightarrow \neg\neg A)$ | Assioma 1 |
| 2. | $\{\neg\neg A\} \vdash \neg\neg A$ | Ipotesi |
| 3. | $\{\neg\neg A\} \vdash \neg\neg\neg A \rightarrow \neg\neg A$ | MP 1,2 |
| 4. | $\{\neg\neg A\} \vdash \neg A \rightarrow \neg\neg A$ | Contrapposizione 3 |
| 5. | $\{\neg\neg A\} \vdash \neg\neg A \rightarrow A$ | Contrapposizione 4 |
| 6. | $\{\neg\neg A\} \vdash A$ | MP 2,5 |
| 7. | $\vdash \neg\neg A \rightarrow A$ | Deduzione 6 |

□

Riguardo la regola di doppia negazione, ricordiamo che essa corrisponde al tipo di negazione che abbiamo preso in considerazione.

Proposizione 8.7 (Consequentia

mirabilis).

$$\vdash (\neg A \rightarrow A) \rightarrow A$$

Dimostrazione.

| | | |
|----|---|--------------------|
| 1. | $\{A\} \vdash A$ | Ipotesi |
| 2. | $\{A \rightarrow \neg A, A\} \vdash A \rightarrow \neg A$ | Ipotesi |
| 3. | $\{A \rightarrow \neg A, A\} \vdash \neg A$ | MP 1,2 |
| 4. | $\{A\} \vdash (A \rightarrow \neg A) \rightarrow \neg A$ | Deduzione 3 |
| 5. | $\{A\} \vdash A \rightarrow \neg(A \rightarrow \neg A)$ | Contrapposizione 4 |
| 6. | $\{A\} \vdash \neg(A \rightarrow \neg A)$ | MP 1,5 |
| 7. | $\vdash A \rightarrow \neg(A \rightarrow \neg A)$ | Deduzione 6 |
| 8. | $\vdash (A \rightarrow \neg A) \rightarrow \neg A$ | Contrapposizione 7 |

□

Da qui, usando il principio dello Pseudo-Scoto, è facile ricavare la giustificazione del “ragionamento per assurdo” (*reductio ad absurdum*), che abbiamo più volte utilizzato. Infatti, se supponendo $\neg A$ si arriva ad una contraddizione, allora, in particolare, da $\neg A$ si può dimostrare A . Usando il teorema deduttivo, varrà allora $\vdash \neg A \rightarrow A$. Per modus ponens dalla consequentia mirabilis, otteniamo $\vdash A$.

Esempio 8.19. *Alcune delle regole*

*appena dimostrate sono anche invertibili.
Così le regole di contrapposizione e
doppia negazione possono essere
formulate nella direzione opposta.*

*La regola ‘invertita’ di
contrapposizione è*

$$(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$$

*e può essere dimostrata nel modo
seguente:*

- | | |
|---|---------------------------|
| 1. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash \neg\neg A$ | <i>Ipotesi</i> |
| 2. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash A$ | <i>Doppia neg. 1</i> |
| 3. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash A \rightarrow B$ | <i>Ipotesi</i> |
| 4. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash B$ | <i>MP 2,3</i> |
| 5. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash \neg B$ | <i>Ipotesi</i> |
| 6. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash \neg B \rightarrow (B \rightarrow \neg\neg B)$ | <i>Pseudo-Scoto</i> |
| 7. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash B \rightarrow \neg\neg B$ | <i>MP 5,6</i> |
| 8. $\{A \rightarrow B, \neg B, \neg\neg A\} \vdash \neg\neg B$ | <i>MP 4,7</i> |
| 9. $\{A \rightarrow B, \neg B\} \vdash \neg\neg A \rightarrow \neg\neg B$ | <i>Deduzione 8</i> |
| 10. $\{A \rightarrow B, \neg B\} \vdash \neg B \rightarrow \neg A$ | <i>Contrapposizione 9</i> |
| 11. $\{A \rightarrow B, \neg B\} \vdash \neg B$ | <i>Ipotesi</i> |
| 12. $\{A \rightarrow B, \neg B\} \vdash \neg A$ | <i>MP 10,11</i> |
| 13. $\{A \rightarrow B\} \vdash \neg B \rightarrow \neg A$ | <i>Deduzione 12</i> |
| 14. $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$ | <i>Deduzione 13</i> |

*Similmente, la regola ‘invertita’ di doppia
negazione è*

$$\vdash A \rightarrow \neg\neg A$$

e può essere dimostrata nel modo seguente:

- | | |
|--|--------------------|
| 1. $\vdash \neg\neg\neg A \rightarrow \neg A$ | <i>Doppia neg.</i> |
| 2. $\vdash (\neg\neg\neg A \rightarrow \neg A) \rightarrow (A \rightarrow \neg\neg A)$ | <i>Assioma 3</i> |
| 3. $\vdash A \rightarrow \neg\neg A$ | <i>MP 1,2</i> |

Come si è visto, nella nostra presentazione del sistema di Hilbert abbiamo fatto uso di due soli connettivi: \neg e \rightarrow . Gli altri possono essere introdotti per definizione usando l’equivalenza logica. Diamo qualche esempio di dimostrazione che coinvolge diversi connettivi.

Esempio 8.20. *Dimostriamo in Hilbert la seguente tautologia:*

$$\vdash A \rightarrow (B \rightarrow (A \wedge B))$$

- | | |
|--|--------------------------------------|
| 1. $\{A, B\} \vdash (A \rightarrow \neg B) \rightarrow (A \rightarrow \neg B)$ | <i>Esempio 8.18</i> |
| 2. $\{A, B\} \vdash A \rightarrow ((A \rightarrow \neg B) \rightarrow \neg B)$ | <i>Scambio 1</i> |
| 3. $\{A, B\} \vdash A$ | <i>Ipotesi</i> |
| 4. $\{A, B\} \vdash (A \rightarrow \neg B) \rightarrow \neg B$ | <i>MP 2,3</i> |
| 5. $\{A, B\} \vdash \neg\neg B \rightarrow \neg(A \rightarrow \neg B)$ | <i>Contrapposizione 4</i> |
| 6. $\{A, B\} \vdash B$ | <i>Ipotesi</i> |
| 7. $\{A, B\} \vdash \neg\neg B$ | <i>Doppia neg. 6</i> |
| 8. $\{A, B\} \vdash \neg(A \rightarrow \neg B)$ | <i>MP 5,7</i> |
| 9. $\{A\} \vdash B \rightarrow \neg(A \rightarrow \neg B)$ | <i>Deduzione 8</i> |
| 10. $\vdash A \rightarrow (B \rightarrow \neg(A \rightarrow \neg B))$ | <i>Deduzione 9</i> |
| 11. $\vdash A \rightarrow (B \rightarrow (A \wedge B))$ | <i>Def. di ‘\wedge’</i> |

Esempio 8.22. Dimostriamo in Hilbert la seguente tautologia:

$$\vdash A \vee (B \vee C) \rightarrow (A \vee B) \vee C$$

| | | |
|-----|--|------------------------------------|
| 1. | $\{\neg A \rightarrow B, \neg B\} \vdash \neg A \rightarrow B$ | <i>Ipotesi</i> |
| 2. | $\{\neg A \rightarrow B, \neg B\} \vdash \neg B \rightarrow \neg\neg A$ | <i>Contrapposizione I</i> |
| 3. | $\{\neg A \rightarrow B, \neg B\} \vdash \neg B$ | <i>Ipotesi</i> |
| 4. | $\{\neg A \rightarrow B, \neg B\} \vdash \neg\neg A$ | <i>MP 2,3</i> |
| 5. | $\{\neg A \rightarrow B, \neg B\} \vdash A$ | <i>Doppia neg. 4</i> |
| 6. | $\{\neg A \rightarrow B\} \vdash \neg B \rightarrow A$ | <i>Deduzione 5</i> |
| 7. | $\vdash (\neg A \rightarrow B) \rightarrow (\neg B \rightarrow A)$ | <i>Deduzione 6</i> |
| 8. | $\vdash A \vee B \rightarrow B \vee A$ | <i>Def. di '\vee'</i> |
| 1. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg(\neg A \rightarrow B)$ | <i>Ipotesi</i> |
| 2. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash B \rightarrow (\neg A \rightarrow B)$ | <i>Axioma 1</i> |
| 3. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg(\neg A \rightarrow B) \rightarrow \neg B$ | <i>Contrap.2</i> |
| 4. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg B$ | <i>MP I,3</i> |
| 5. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash A \rightarrow (\neg A \rightarrow B)$ | <i>Pseudo-Scoto</i> |
| 6. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg(\neg A \rightarrow B) \rightarrow \neg A$ | <i>Contrap.5</i> |
| 7. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg A$ | <i>MP I,7</i> |
| 8. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg A \rightarrow (\neg B \rightarrow C)$ | <i>Ipotesi</i> |
| 9. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash \neg B \rightarrow C$ | <i>MP 7,8</i> |
| 10. | $\{\neg A \rightarrow (\neg B \rightarrow C), \neg(\neg A \rightarrow B)\} \vdash C$ | <i>MP 4,9</i> |
| 11. | $\{\neg A \rightarrow (\neg B \rightarrow C)\} \vdash \neg(\neg A \rightarrow B) \rightarrow C$ | <i>Deduz.10</i> |
| 12. | $\vdash (\neg A \rightarrow (\neg B \rightarrow C)) \rightarrow (\neg(\neg A \rightarrow B) \rightarrow C)$ | <i>Deduz.11</i> |
| 13. | $\vdash A \vee (B \vee C) \rightarrow (A \vee B) \vee C$ | <i>Def. di '\vee'</i> |

Analogamente a quanto fatto per il metodo dei tableau proposizionali e per il sistema di Gentzen, diamo il teorema di completezza e di correttezza.

Teorema 8.3 (Correttezza e completezza del sistema H). Una formula è valida se e soltanto se è

dimostrabile nel sistema di Hilbert.

Dimostrazione. La dimostrazione di correttezza si conduce notando che gli assiomi sono validi perchè è possibile costruire tableau chiusi semantici per le loro negazioni (il lettore può farlo per esercizio); se la regola di inferenza (*Modus Ponens*) non fosse corretta, l'insieme di formule $\{A, A \rightarrow B, B\}$ sarebbe tale che sia A che $A \rightarrow B$ sono valide, ma non lo è B . Allora esisterebbe un'interpretazione v tale che $v(B) = F$. Dalla validità di A e di $A \rightarrow B$ segue che $v(A) = v(A \rightarrow B) = V$ per ogni interpretazione; e da qui $v(B) = V$, in contraddizione con quanto sopra posto.

Per quanto riguarda la completezza del sistema di Hilbert, si può dimostrare che ogni dimostrazione nel sistema di Gentzen può essere meccanicamente trasformata in una dimostrazione nel sistema di

Hilbert; e sappiamo inoltre che ogni formula valida può essere verificata nel sistema di Gentzen (per i dettagli della dimostrazione, che è laboriosa, ma non difficile, rimandiamo a [6]).

□

Osserviamo come questo risultato (come d'altronde l'analogo per il sistema di Gentzen) ponga una corrispondenza perfetta tra la *sintassi* e la *semantica* del linguaggio, facendo coincidere la *dimostrabilità* con la *validità*:

$\vdash P$ se e soltanto se $|= P$.

Teorema 8.4 (Coerenza). *I sistemi G ed H sono coerenti, cioè non è possibile provare in essi P e $\neg P$.*

Dimostrazione. Si tratta ancora di una semplice conseguenza del teorema di

correttezza, per il quale ciò che è dimostrabile in uno di questi sistemi è anche valido per le tavole di verità; e questo non può accadere contemporaneamente per P e $\neg P$.

□

Esercizi svolti

Esercizio 8.1. *Quali dei seguenti enunciati è vero?*

1. *Roma è in Italia e $1 + 1 = 3$;*
2. *Roma è in Spagna e $1 + 1 = 2$;*
3. *Roma è in Spagna e $1 + 1 = 3$;*
4. *Roma è in Italia e $1 + 1 = 2$.*

Come cambierebbero le risposte sostituendo alla congiunzione la disgiunzione (o)? E l'implicazione (se... allora...)? E la doppia implicazione (se e

soltanto se)?

Soluzione: Nel caso della congiunzione, l'unico enunciato vero è l'ultimo, poichè una congiunzione è vera se e soltanto se entrambi i congiunti sono veri. Nel caso della disgiunzione, l'unico enunciato falso è il terzo, poichè una disgiunzione è falsa se e soltanto se entrambi i disgiunti sono falsi. Nel caso della implicazione, l'unico enunciato falso è il primo, poichè un'implicazione è falsa se e soltanto se la premessa è vera e la conseguenza è falsa. Infine, nel caso della doppia implicazione, gli enunciati veri sono gli ultimi due, poichè una doppia implicazione è vera se e soltanto i due enunciati hanno lo stesso valore di verità.

Esercizio 8.2. Siano $p = \text{"Luigi è alto"}$ e $q = \text{"Luigi è bello"}$. Scrivere ciascuna delle seguenti frasi in forma di enunciato

usando p e q :

1. *Luigi è alto e bello;*
2. *Luigi è alto ma non bello;*
3. *è falso che Luigi è basso o bello;*
4. *Luigi non è bello né alto;*
5. *Luigi è alto, oppure è basso e bello;*
6. *non è vero che Luigi è basso o brutto.*

Soluzione:

1. $p \wedge q;$
2. $p \wedge \neg q;$
3. $\neg(\neg p \vee q);$
4. $\neg p \wedge \neg q;$
5. $p \vee (\neg p \wedge q);$
6. $\neg(\neg p \vee \neg q).$

Esercizio 8.3. *Costruire la tavola di verità di ciascuno dei seguenti enunciati:*

1. $\neg p \wedge q$;
2. $\neg(p \rightarrow \neg q)$;
3. $(p \wedge q) \rightarrow (p \vee q)$.

Soluzione: La tavola per il primo enunciato è:

| p | q | $\neg p$ | $\neg p \wedge q$ |
|-----|-----|----------|-------------------|
| F | F | T | F |
| F | T | T | T |
| T | T | F | F |

La tavola per il secondo enunciato è:

| p | q | $\neg q$ | $p \rightarrow \neg q$ | $\neg(p \rightarrow \neg q)$ |
|-----|-----|----------|------------------------|------------------------------|
| F | F | T | T | F |
| F | T | F | T | F |
| T | F | T | T | F |
| T | T | F | F | T |

La tavola per il terzo enunciato è:

| p | q | $p \wedge q$ | $p \vee q$ | $(p \wedge q) \rightarrow (p \vee q)$ |
|-----|-----|--------------|------------|---------------------------------------|
| F | F | F | F | T |
| F | T | F | T | T |
| T | F | F | T | T |
| T | T | T | T | T |

Esercizio 8.4. Verificare, mediante le tavole di verità, le seguenti equivalenze logiche:

1. $\neg(p \vee q) \equiv \neg p \wedge \neg q$ (Legge di De Morgan);
2. $\neg(p \rightarrow q) \equiv p \wedge \neg q$;
3. $\neg\neg p \equiv p$.

Soluzione: Si comparino le colonne evidenziate in grassetto (corrispondenti alle formule equivalenti):

| p | q | $p \vee q$ | $\neg(p \vee q)$ | $\neg p$ | $\neg q$ | $\neg p \wedge \neg q$ |
|-----|-----|------------|------------------|----------|----------|------------------------|
| F | F | F | T | T | T | T |
| F | T | T | F | T | F | F |
| T | F | T | F | F | T | F |
| T | T | T | F | F | F | F |

| p | q | $p \rightarrow q$ | $\neg(p \rightarrow q)$ | $\neg q$ | $p \wedge \neg q$ |
|-----|-----|-------------------|-------------------------|----------|-------------------|
| F | F | T | F | T | F |
| F | T | T | F | F | F |
| T | F | F | T | T | T |
| T | T | T | F | F | F |

| p | $\neg p$ | $\neg \neg p$ |
|-----|----------|---------------|
| F | T | F |
| T | F | T |

Esercizio 8.5. Usare i risultati dell'[Esercizio 8.4](#) per semplificare i seguenti enunciati:

1. $\neg(\neg p \rightarrow q)$;
2. $\neg(p \wedge \neg q)$;
3. $\neg(\neg p \wedge \neg q)$.

Soluzione:

1. Per la seconda equivalenza dell'[Esercizio 8.4](#), $\neg(\neg p \rightarrow q) \equiv \neg p \wedge \neg q$; per la prima equivalenza dell'[Esercizio 8.4](#), $\neg p \wedge \neg q \equiv \neg(p \vee q)$. Quest'ultimo enunciato è più semplice di quello dato, in quanto ha un connettivo in meno.
2. Per la seconda equivalenza

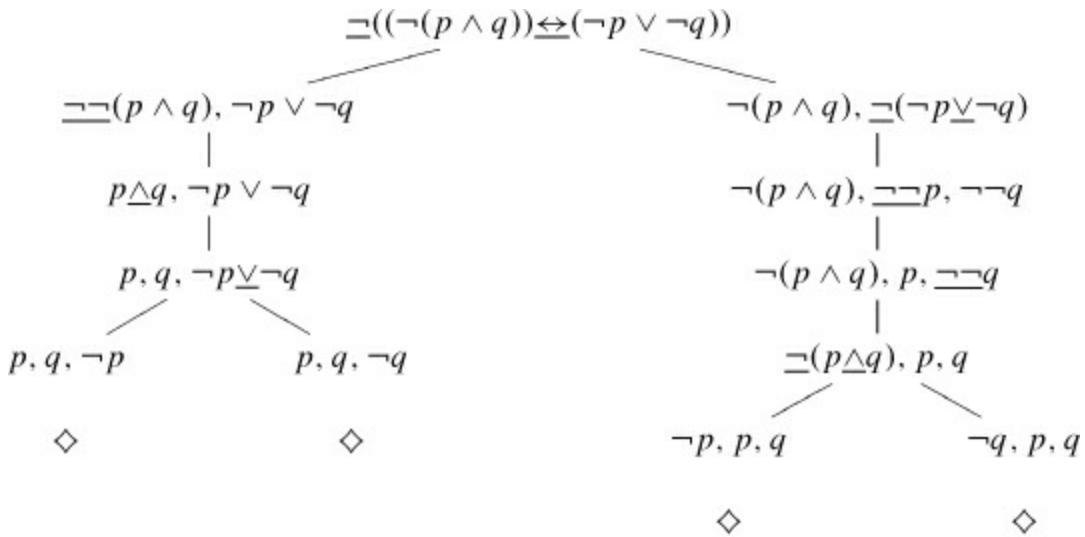
dell'[Esercizio 8.4](#), $\neg(p \wedge \neg q) \equiv \neg\neg(p \rightarrow q)$; per la terza equivalenza dell'[Esercizio 8.4](#), $\neg\neg(p \rightarrow q) \equiv p \rightarrow q$.

3. Per la prima equivalenza dell'[Esercizio 8.4](#), $\neg(\neg p \wedge \neg q) \equiv \neg\neg(p \vee q)$; per la terza equivalenza dell'[Esercizio 8.4](#), $\neg\neg(p \vee q) \equiv p \vee q$.

Esercizio 8.6. Verificare usando i tableau la seguente legge di De Morgan:

$$\neg(p \wedge q) \equiv \neg p \vee \neg q$$

Soluzione: Costruiremo il tableau per la negata della formula $(\neg(p \wedge q)) \leftrightarrow (\neg p \vee \neg q)$; siccome risulterà chiuso, avremo che $(\neg(p \wedge q)) \leftrightarrow (\neg p \vee \neg q)$ è una tautologia e pertanto l'equivalenza data vale.



Spesso, per scrivere l'albero in maniera più compatta, accorperemo regole successive dello stesso tipo; per esempio, nell'albero precedente, scriveremo

$$\begin{array}{c}
\neg(p \wedge q), \underline{\neg}\neg p, \underline{\neg}\neg q \\
| \\
\neg(p \wedge q), p, q
\end{array}$$

invece di

$$\begin{array}{c}
\neg(p \wedge q), \underline{\neg}\neg p, \neg\neg q \\
| \\
\neg(p \wedge q), p, \underline{\neg}\neg q \\
| \\
\neg(p \wedge q), p, q
\end{array}$$

Si noti inoltre che, in generale, il tableau

risulta più piccolo se si applicano prima le α -regole e poi le β -regole. Ad esempio, se nel costruire il tableau per $\neg((\neg(p \wedge q)) \leftrightarrow (\neg p \vee \neg q))$ avessimo sempre scelto (quando possibile) l'applicazione di β -regole, avremmo ottenuto un tableau con 17 nodi contro i 12 del tableau visto in precedenza.

Esercizio 8.7. *Usando il metodo dei tableau, dire se il seguente enunciato è una tautologia:*

$$A \rightarrow (A \rightarrow \neg A)$$

Soluzione: Proviamo a costruire il tableau della negata:

$$\begin{array}{c}
 \neg(A \rightarrow (A \rightarrow \neg A)) \\
 | \\
 A, \neg(A \rightarrow \neg A) \\
 | \\
 A, A, \neg\neg A \\
 | \\
 A, A, A
 \end{array}$$

Questo tableau non è chiuso e, pertanto, la formula data non è una tautologia. Le foglie aperte di un tableau suggeriscono un'interpretazione che falsifica la negata della formula alla radice: nel nostro caso, assegnando ad A valore T si ottiene $T \rightarrow (T \rightarrow F)$ il cui valore è F .

Esercizio 8.8. *Usando il metodo dei tableau e l'induzione, dimostrare che l'enunciato*

$$F_n = \begin{cases} A \rightarrow (A \rightarrow A) & \text{se } n = 0 \\ A \rightarrow F_{n-1} & \text{altrimenti} \end{cases}$$

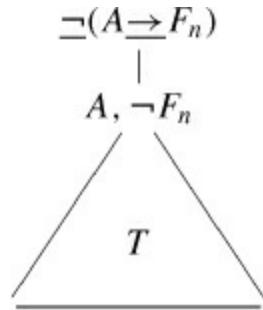
è valido per ogni $n \in \mathbb{N}$.

Soluzione: Dimostriamo, per induzione su n , che il tableau di $\neg F_n$ è chiuso. Il passo base è per $n = 0$:

$$\begin{array}{c} \neg(A \rightarrow (A \rightarrow A)) \\ | \\ A, \neg(A \rightarrow A) \\ | \\ A, A, \neg A \end{array}$$

◊

Per quanto riguarda il passo induttivo, assumiamo che il tableau di $\neg F_n$ sia chiuso e dimostriamo che il tableau di $\neg F_{n+1}$ è chiuso. Il tableau di $\neg F_{n+1}$ sarà della forma:



dove il tableau T è ottenuto dal tableau di $\neg F_n$ aggiungendo ad ogni nodo l'enunciato

A. Poichè il tableau di $\neg F_n$ ha in ogni foglia una coppia di enunciati contraddittori, anche T ha questa proprietà e quindi è chiuso. Pertanto, F_{n+1} è una tautologia, essendo il tableau della sua negata chiuso.

Esercizio 8.9. Svolgere l'[Esercizio 8.6](#) usando il metodo di Gentzen.

Soluzione:

$$\begin{array}{ccc}
 \frac{\checkmark}{\overbrace{p, \neg p, \neg q}} & \frac{\checkmark}{\overbrace{q, \neg p, \neg q}} & \\
 \frac{p \Delta q, \neg p, \neg q}{p \wedge q, \neg p \vee \neg q} & \frac{\checkmark}{\overbrace{\neg \neg p, \neg p, \neg q}} & \frac{\checkmark}{\overbrace{q, \neg p, \neg q}} \\
 \frac{\neg \neg(p \wedge q), \neg p \vee \neg q}{(\neg(p \wedge q)) \rightarrow (\neg p \vee \neg q)} & \frac{\neg \neg(\neg p \vee \neg q), \neg p, \neg q}{\neg(\neg p \vee \neg q), \neg(p \Delta q)} & \\
 & \frac{\neg(\neg p \vee \neg q), \neg(p \Delta q)}{(\neg p \vee \neg q) \rightarrow (\neg(p \wedge q))} & \\
 & & \frac{(\neg(p \wedge q)) \leftrightarrow (\neg p \vee \neg q)}{(\neg(p \wedge q)) \leftrightarrow (\neg p \vee \neg q)}
 \end{array}$$

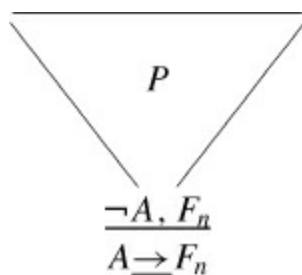
Si noti che tutti gli insiemi di enunciati segnati \checkmark con sono assiomi nel sistema di Gentzen, perchè contengono un enunciato e la sua negazione.

Esercizio 8.10. Si risolva con l'induzione ed il metodo di Gentzen l'Esercizio 8.8.

Soluzione: Dimostriamo per induzione su n che F_n ha una prova nel sistema di Gentzen. Il passo base è per $n = 0$:

$$\frac{\begin{array}{c} \neg A, \neg A, A \\ \hline \neg A, A \underline{\rightarrow} A \end{array}}{A \underline{\rightarrow} (A \rightarrow A)}$$

Per il passo induttivo, osserviamo che la seguente è una prova di F_{n+1} nel sistema di Gentzen:



dove P è la prova per F_n (che esiste per ipotesi induttiva) in cui si aggiunge $\neg A$ ad

ogni insieme di enunciati. Quello che si ottiene è ancora una prova nel sistema di Gentzen, poichè tutti gli insiemi di formule senza premesse in questa prova sono assiomi (ogni assioma nella prova per F_n resta un assioma se gli uniamo $\neg A$), e dimostra F_{n+1} .

Esercizio 8.11. *Si dimostri, usando il metodo di Hilbert, che $\vdash A \rightarrow F$ per ogni enunciato F tale che $\vdash F$.*

Soluzione: Poichè $\vdash F$, abbiamo anche che $A \vdash F$; usando il teorema deduttivo, concludiamo che $\vdash A \rightarrow F$.

Esercizio 8.12. *Si dimostri, usando il metodo di Hilbert, la seguente generalizzazione della regola di scambio delle premesse:*

$$\frac{U \vdash A \rightarrow (B \rightarrow (C \rightarrow D))}{U \vdash A \rightarrow (C \rightarrow (B \rightarrow D))}$$

Soluzione: Per il teorema deduttivo, $U \vdash A \rightarrow (B \rightarrow (C \rightarrow D))$ vale se e solo se $U, A \vdash B \rightarrow (C \rightarrow D)$. Per la regola di scambio delle premesse, questo vale se e solo se $U, A \vdash C \rightarrow (B \rightarrow D)$ che, nuovamente per il teorema deduttivo, equivale a $U \vdash A \rightarrow (C \rightarrow (B \rightarrow D))$.

Esercizio 8.13. *Si dimostri, usando il metodo di Hilbert, che $\vdash (A \rightarrow (B \rightarrow C)) \rightarrow (A \rightarrow ((A \rightarrow B) \rightarrow C))$.*

Soluzione: Si conclude facilmente partendo dal secondo assioma, usando la regola di scambio delle premesse vista nell'Esercizio 8.12:

$$\begin{aligned} &\vdash (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)) \\ &\vdash (A \rightarrow (B \rightarrow C)) \rightarrow (A \rightarrow ((A \rightarrow B) \rightarrow C)) \end{aligned}$$

Esercizio 8.14. *Si dimostri, usando il*

metodo di Hilbert, che $\vdash \neg A \rightarrow \neg\neg(A \rightarrow \neg B)$.

Soluzione: Partiamo da un’istanza del primo assioma e applichiamo in sequenza le regole di contrapposizione e di doppia negazione generalizzate dell’Esercizio 8.28:

$$\begin{aligned} &\vdash \neg A \rightarrow (\neg\neg B \rightarrow \neg A) \\ &\vdash \neg A \rightarrow (A \rightarrow \neg B) \\ &\vdash \neg A \rightarrow \neg\neg(A \rightarrow \neg B) \end{aligned}$$

Esercizio 8.15. *Si dimostri, usando il metodo di Hilbert, che* $\vdash (A \rightarrow (B \rightarrow C)) \vdash ((A \wedge B) \rightarrow C)$.

$$\begin{array}{ll} A \rightarrow (B \rightarrow C) \vdash A \rightarrow (B \rightarrow C) & \text{Ip.} \\ A \rightarrow (B \rightarrow C) \vdash A \rightarrow (\neg C \rightarrow \neg B) & \text{Es. 8.28} \\ A \rightarrow (B \rightarrow C) \vdash \neg C \rightarrow (A \rightarrow \neg B) & \text{Scambio Prem.} \\ A \rightarrow (B \rightarrow C) \vdash \neg C \rightarrow \neg\neg(A \rightarrow \neg B) & \text{Es. 8.28} \\ A \rightarrow (B \rightarrow C) \vdash \neg(A \rightarrow \neg B) \rightarrow C & \text{Contrapp.} \\ & \vdash (A \rightarrow (B \rightarrow C)) \rightarrow (\neg(A \rightarrow \neg B) \rightarrow C) \text{ Deduzione} \end{array}$$

Soluzione: Anzitutto, riscriviamo l’enunciato dato in termini di implicazioni e negazioni; in particolare, $A \wedge B \equiv \neg(A \rightarrow \neg B)$.

Esercizio 8.16. Si dimostri, usando il metodo di Hilbert, che $\vdash ((A \vee B) \rightarrow C) \vdash ((A \rightarrow C) \vee (B \rightarrow C))$.

Soluzione: Anzitutto, riscriviamo l'enunciato dato in termini di implicazioni e negazioni; in particolare, $A \vee B \equiv \neg A \rightarrow B$ e $(A \rightarrow C) \vee (B \rightarrow C) \equiv \neg(A \rightarrow C) \rightarrow (B \rightarrow C)$.

| | |
|--|-----------|
| $(\neg A \rightarrow B) \rightarrow C, \neg(A \rightarrow C), B, \neg A \vdash B$ | I.p. |
| $(\neg A \rightarrow B) \rightarrow C, \neg(A \rightarrow C), B \vdash \neg A \rightarrow B$ | Deduzione |
| $(\neg A \rightarrow B) \rightarrow C, \neg(A \rightarrow C), B \vdash \neg(\neg A \rightarrow B) \rightarrow C$ | I.p. |
| $(\neg A \rightarrow B) \rightarrow C, \neg(A \rightarrow C), B \vdash C$ | M.P. 2,3 |
| $(\neg A \rightarrow B) \rightarrow C, \neg(A \rightarrow C) \vdash B \rightarrow C$ | Deduzione |
| $(\neg A \rightarrow B) \rightarrow C \vdash \neg(A \rightarrow C) \rightarrow (B \rightarrow C)$ | Deduzione |
| $(\neg A \rightarrow B) \rightarrow C \vdash \neg((\neg A \rightarrow B) \rightarrow C) \rightarrow$ | Deduzione |
| $(\neg A \rightarrow C) \rightarrow (B \rightarrow C))$ | |

Esercizio 8.17. Si dimostri, usando il metodo di Hilbert, che $\vdash (B \wedge \neg A) \rightarrow (\neg C \rightarrow \neg(A \vee C))$.

Soluzione: Anzitutto, riscriviamo l'enunciato dato in termini di implicazioni e negazioni; in particolare, $B \wedge \neg A \equiv \neg(B \rightarrow A)$ e $A \vee C \equiv \neg A \rightarrow C$.

| | |
|--|-----------|
| $\neg(B \rightarrow A), \neg A \rightarrow C \vdash A \rightarrow (B \rightarrow A)$ | Ax.1 |
| $\neg(B \rightarrow A), \neg A \rightarrow C \vdash \neg(B \rightarrow A) \rightarrow \neg A$ | Contrapp. |
| $\neg(B \rightarrow A), \neg A \rightarrow C \vdash \neg(B \rightarrow A)$ | Ip. |
| $\neg(B \rightarrow A), \neg A \rightarrow C \vdash \neg A$ | M.P.2,3 |
| $\neg(B \rightarrow A), \neg A \rightarrow C \vdash \neg A \rightarrow C$ | Ip. |
| $\neg(B \rightarrow A), \neg A \rightarrow C \vdash C$ | M.P.4,5 |
| $\neg(B \rightarrow A) \quad \vdash (\neg A \rightarrow C) \rightarrow C$ | Deduzione |
| $\neg(B \rightarrow A) \quad \vdash \neg C \rightarrow \neg(\neg A \rightarrow C)$ | Contrapp. |
| $\neg(B \rightarrow A) \quad \vdash \neg(B \rightarrow A) \rightarrow (\neg C \rightarrow \neg(\neg A \rightarrow C))$ | Deduzione |

Esercizio 8.18. Si dimostri per induzione e usando il metodo di Hilbert che gli enunciati $\{F_n\}_{n \in \mathbb{N}_0}$ dell'[Esercizio 8.27](#) sono tutti tautologici.

Soluzione: Per il passo base, dobbiamo dimostrare $A \rightarrow (B \rightarrow (A \rightarrow B))$ nel sistema di Hilbert: questo può essere fatto osservando che $\vdash B \rightarrow (A \rightarrow B)$ è un'istanza del primo assioma e usando l'Esercizio 8.11. Per il passo induttivo, sappiamo che $\vdash F_n$; usando due volte l'Esercizio 8.11, otteniamo che $\vdash A \rightarrow (B \rightarrow F_n)$, cioè $\vdash F_{n+1}$.

Esercizi da svolgere

Esercizio 8.19. Formalizzare e determinare tramite tavole di verità il valore di ciascuno dei seguenti enunciati:

1. se $3 + 2 = 7$, allora $4 + 4 = 8$;
2. non è vero che: $2 + 2 = 5$ sse $4 + 4 = 10$;
3. Parigi è in Inghilterra o Londra è in Francia;
4. non è vero che $1 + 1 = 3$ o $2 + 1 = 3$;
5. è falso che se Parigi è in Inghilterra allora Londra è in Francia.

Esercizio 8.20. Costruire la tavola di verità di ciascuno dei seguenti enunciati:

1. $\neg(p \wedge q) \vee \neg(q \leftrightarrow p)$;
2. $(p \rightarrow q) \vee \neg(p \leftrightarrow \neg q)$;
3. $(p \rightarrow (\neg q \vee r)) \wedge \neg(q \vee (p \leftrightarrow \neg r))$.

Esercizio 8.21. Verificare, mediante le tavole di verità, le seguenti equivalenze

logiche:

1. $\neg(p \wedge q) \equiv \neg p \vee \neg q$ (*Legge di De Morgan*);
2. $\neg(p \leftrightarrow q) \equiv p \leftrightarrow \neg q \equiv \neg p \leftrightarrow q.$

Esercizio 8.22. *Usare i risultati degli Esercizi 8.4 e 8.21 per semplificare i seguenti enunciati:*

1. $\neg(p \vee \neg q);$
2. $\neg(\neg p \leftrightarrow q);$
3. $\neg(\neg p \rightarrow \neg q).$

Esercizio 8.23. *Semplificare ciascuna delle frasi seguenti:*

1. *non è vero che se le rose sono rosse allora le violette sono blu;*
2. *non è vero che fa freddo e piove;*
3. *non è vero che Luigi è basso o bello;*

4. *non è vero che non fa freddo o piove;*
5. *non è vero che se piove allora fa freddo;*
6. *non è vero che: le rose sono rosse se e solo se le violette sono blu.*

Esercizio 8.24. Verificare tutte le equivalenze logiche viste negli [Esercizi 8.4](#) e [8.21](#) usando sia il metodo dei tableau che il metodo di Gentzen.

Esercizio 8.25. Usando sia il metodo dei tableau che il metodo di Gentzen, si dimostrino le seguenti equivalenze:

1. $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r);$
2. $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r);$
3. $p \rightarrow (q \wedge r) \equiv (p \rightarrow q) \wedge (p \rightarrow r).$

Esercizio 8.26. Usando sia il metodo dei tableau che il metodo di Gentzen, verificare se i seguenti enunciati sono

tautologie:

1. $((A \vee B) \rightarrow C) \rightarrow ((A \rightarrow C) \vee (B \rightarrow C));$
2. $((A \rightarrow C) \wedge (B \rightarrow C)) \rightarrow ((A \wedge B) \rightarrow C);$
3. $\neg(A \rightarrow B) \rightarrow ((A \rightarrow C) \rightarrow \neg B);$
4. $(A \rightarrow B) \vee (B \rightarrow C) \vee (C \rightarrow A);$
5. $(A \rightarrow B) \rightarrow ((A \rightarrow C) \rightarrow (A \rightarrow (B \wedge C)));$
6. $((A \rightarrow C) \wedge (B \rightarrow C)) \rightarrow ((A \vee B) \rightarrow C).$

Esercizio 8.27. Si dica quali dei seguenti enunciati sono validi:

$$F_n = \begin{cases} A \rightarrow B & \text{se } n = 0 \\ A \rightarrow (B \rightarrow F_{n-1}) & \text{altrimenti} \end{cases}$$

$$G_n = \begin{cases} A \rightarrow (B \rightarrow A) & \text{se } n = 0 \\ A \rightarrow (B \rightarrow G_{n-1}) & \text{altrimenti} \end{cases}$$

Si proceda per induzione, usando sia il metodo dei tableau che il metodo di Gentzen.

Esercizio 8.28. Si dimostrino, usando il

metodo di Hilbert, le seguenti generalizzazioni delle regole di contrapposizione, doppia negazione e transitività:

$$\frac{U \vdash A \rightarrow (B \rightarrow C)}{U \vdash A \rightarrow (\neg C \rightarrow \neg B)} \quad \frac{U \vdash A \rightarrow \neg\neg B}{U \vdash A \rightarrow B}$$

$$\frac{\begin{array}{c} U \vdash A \rightarrow (B \rightarrow C) \quad U \vdash A \rightarrow (C \rightarrow D) \end{array}}{U \vdash A \rightarrow (B \rightarrow D)}$$

Esercizio 8.29. *Si dimostri, usando il metodo di Hilbert, che $\vdash (A \rightarrow B) \rightarrow (A \rightarrow (B \rightarrow (A \rightarrow B)))$.*

Esercizio 8.30. *Si dimostri, usando il metodo di Hilbert, che $\vdash (A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$.*

Esercizio 8.31. *Si dimostri, usando il metodo di Hilbert, che $\vdash A \rightarrow (A \vee (B \wedge C))$.*

Esercizio 8.32. *Si dimostri, usando il*

metodo di Hilbert, che $\vdash (A \rightarrow ((B \rightarrow C) \rightarrow D)) \rightarrow (\neg(A \rightarrow D) \rightarrow (A \wedge (B \wedge \neg C)))$.

Esercizio 8.33. *Si dimostri, usando il metodo di Hilbert, che $\vdash A \rightarrow (\neg A \rightarrow B)$.*

Esercizio 8.34. *Si dimostri, usando il metodo di Hilbert, che $\vdash (A \rightarrow \neg A) \rightarrow \neg A$.*

Esercizio 8.35. *Si dimostri, usando il metodo di Hilbert, che $\vdash A \rightarrow A \vee B$.*

Esercizio 8.36. *Si dimostri, usando il metodo di Hilbert, che $\vdash (A \rightarrow B) \rightarrow ((C \vee A) \rightarrow (C \vee B))$.*

Esercizio 8.37. *Si dimostri, usando il metodo di Hilbert, che $\vdash (A \vee B) \vee C \rightarrow A \vee (B \vee C)$.*

Esercizio 8.38. *Si dia una dimostrazione più semplice della Proposizione 8.5.*

Esercizio 8.39. Si dimostri per induzione e usando il metodo di Hilbert che gli enunciati $\{F_n\}_{n \in \mathbb{N}}$ dell'[Esercizio 8.8](#) e $\{G_n\}_{n \in \mathbb{N}}$ dell'[Esercizio 8.27](#) sono tutti tautologici.

¹ Evert Willem Beth (1908-1964) fu un filosofo e logico olandese che si occupò principalmente dei fondamenti della matematica. Il suo nome è spesso ricordato in relazione ai tableau semantici, un metodo messo a punto indipendentemente da Beth (1955), Hintikka (1955), Schutte (1956), e più tardi sviluppato da Smullyan (1968). I principali contributi di Beth alla logica furono inoltre il teorema di definibilità e i modelli di Beth. Nell'ultimo periodo della sua vita, Beth si dedicò a settori collegati alla logica, come lo studio del linguaggio, la dimostrazione di teoremi, la matematica euristica e i metodi per la traduzione nei linguaggi naturali.

² *Nota storica:* Il logico e matematico tedesco Gerhard Gentzen (1909-1945) è stato allievo di Bernays, Courant, Landau, Weyl e di Hilbert. La sua tesi di dottorato (1933), con Weil, era dedicata ai fondamenti della matematica; a Gottinga nel 1934 fu assistente di Hilbert e sviluppò ricerche in logica, fondamenti della matematica e teoria della dimostrazione. Ottenuta l'abilitazione alla docenza universitaria, dal 1943 insegnò a Praga dove morì nel 1945 in un campo di prigione.

3 Nota storica: Il tedesco David Hilbert (1862-1943) è stato una delle figure più significative della matematica tra il XIX e il XX secolo. Conseguito il dottorato con Lindemann, insegnò all'Università di Königsberg dal 1886 al 1895 per passare poi all'Università di Göttingen, dove restò fino alla fine della carriera. La scuola di Hilbert, ad esempio con Bernays, ha apportato un contributo essenziale alla logica contemporanea con la teoria della dimostrazione.

Calcolo dei predicati

9.1 Logica dei predicati: calcolo dei quantificatori

“La logica dei predicati si presenta come un’estensione dell’algebra delle proposizioni. Essa comprende tutta l’algebra delle proposizioni: cioè le proposizioni elementari, considerate come grandezze che assumono uno dei due valori di verità V o F, tutte le operazioni dell’algebra delle proposizioni e, di conseguenza, tutte le sue formule. In più, però, la logica dei predicati introduce,

nello studio delle proposizioni, attributi di oggetti. In tale logica le proposizioni vengono analizzate in soggetto e predicato”.

P. S. Novikov

9.1.1 Dall’alfabeto alle formule predicative

Come abbiamo precedentemente osservato, la logica che abbiamo finora esposto, detta logica degli enunciati, si occupa degli enunciati atomici intesi come blocchi unici e ne studia le combinazioni, senza esaminare la loro “struttura interna”. In altre parole, nell’ambito della logica degli enunciati le due frasi:

Parigi è in Francia
Tutti i quadrati hanno quattro lati

non presentano alcuna differenza sostanziale. Si tratta infatti di due enunciati (peraltro entrambi veri nel nostro mondo), e dicendo ciò non si fa riferimento alla loro ben diversa struttura: il primo di essi è evidentemente riferito ad un *singolo* soggetto (Parigi), al quale viene attribuita una certa proprietà (quella di trovarsi in Francia); la seconda frase, invece, coinvolge *tutti* i quadrati (quindi ha molti soggetti) e ad essi (ovvero: a ciascun quadrato) riferisce la proprietà di avere quattro lati.

Condurremo ora un'analisi più dettagliata di un'espressione formale, che potrà essere talvolta anche più generale di un enunciato: l'insieme dei concetti e dei procedimenti che dovremmo trattare viene detto *logica dei predicati* (o *calcolo dei predicati*). Esso contiene come parte propria la logica degli enunciati.

Riprendiamo l'introduzione che era stata proposta per la logica degli enunciati; l'alfabeto era costituito da lettere maiuscole per indicare enunciati, da connettivi e da alcuni segni come le parentesi o la virgola. Tale alfabeto non è sufficiente nel caso della logica dei predicati: infatti all'interno delle formule elementari (*atomiche*) del nostro linguaggio andremo anche ad analizzare gli “individui” dei quali si predica qualcosa ed i “predicati”. Esaminando le due frasi sopra riportate, notiamo, a proposito del soggetto, che, nel primo caso, si tratta di un individuo ben preciso e con un nome (Parigi), mentre, nel secondo caso, si tratta di individui generici di una classe (i quadrati), della quale specifichiamo che vanno considerati tutti gli elementi.

Avremmo potuto considerare anche una situazione intermedia, quella, cioè, nella quale gli individui considerati non sono

così generici e nemmeno fissati, ma soltanto parzialmente determinati in base ad altri individui:

I quadrati dei numeri reali sono positivi

In questo caso prendiamo in esame soltanto quegli elementi nella classe dei reali che si ottengono da un altro elemento mediante elevamento al quadrato, sono, cioè, nell'immagine di una funzione definita sulla classe. I nostri individui, ossia i *termini* del nostro linguaggi saranno dunque *costanti*, *variabili* o definiti da altri termini mediante *funzioni* (ad uno o più posti).

Per quanto riguarda i predicati, possiamo osservare che, anche se nella lingua parlata potrebbe sembrare innaturale, ogni predicato può essere ridotto ad una copula seguita da una

proprietà ad uno o più posti, da pensare come una relazione tra individui. Nel primo e nel terzo esempio la cosa è evidente; il secondo potrebbe essere riformulato così:

Tutti i quadrati sono a quattro lati

Come esempio di predicato a più posti consideriamo

Luigi e Giacomo sono fratelli

dove “essere fratelli” è una proprietà che si deve predicare di due individui.

L’alfabeto di un linguaggio per la logica dei predicati si compone di:

1. simboli per variabili, denotati da x, y, z, \dots (eventualmente con indici);

2. simboli per costanti, denotati da a, b, c, \dots (eventualmente con indici);
3. simboli funzionali a 1, 2,... posti (si dice di *arità* 1, 2,...), denotati da f, g, h, \dots (eventualmente con indici);
4. simboli predicativi a 1, 2,... posti, denotati da $P(-), Q(-, -), \dots$ (eventualmente con indici);
5. i connettivi, $\neg, \rightarrow, \wedge, \vee, \leftrightarrow;$
6. i quantificatori $\forall, \exists;$
7. parentesi e virgolette.

I *termini* di un linguaggio per la logica dei predicati sono definiti induttivamente rispetto alla presenza in essi di simboli funzionali:

1. variabili e costanti sono termini;
2. se t_1, \dots, t_n sono termini ed f ha n posti, $f(t_1, \dots, t_n)$ è un termine.

Le *formule* sono espressioni del linguaggio definite induttivamente:

1. se P ha n posti e t_1, \dots, t_n sono termini, $P(t_1, \dots, t_n)$ è una formula (*atomica*);
2. se A e B sono formule, lo sono anche A , $A \rightarrow B$, $A \wedge B$, $A \vee B$ e $A \leftrightarrow B$;
3. se A è una formula, lo sono anche $\forall x A$ e $\exists x A$, dove x è una variabile.

Come è accaduto nella logica degli enunciati, spesso considereremo per semplicità soltanto una parte dell’alfabeto. Talvolta nella scrittura si ometteranno virgolette e parentesi, quando ciò non generi confusione.

9.1.2 I quantificatori

Nella logica dei predicati, spesso per delimitare l’estensione di un soggetto, avremo bisogno di “quantificarlo”;

considereremo quindi frasi del tipo:

*Esiste (almeno) un oggetto x che verifica
la proprietà P*

*Per ogni oggetto y è verificata la
proprietà Q*

La formalizzazione della prima frase sopra riportata necessita di un *quantificatore esistenziale*, \exists , che garantisca l'esistenza di almeno un elemento della classe considerata, che verifichi una proprietà data; la seconda, di un *quantificatore universale*, \forall , che garantisca il rispetto della proprietà data da parte di tutti gli elementi di una classe considerata.

In base a questa intuizione possiamo introdurre il concetto di verità nella logica dei predicati. La verità di una formula complessa non si potrà ottenere, come per

gli enunciati, assegnando dei valori di verità alle componenti atomiche e facendo poi il calcolo in base ai connettivi, usando le tavole di verità. Questa volta le formule atomiche potranno variare il loro valore in base alla struttura interna. Ad esempio, consideriamo la seguente formula A :

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

Essa afferma che *per ogni x esiste un y tale che la coppia (x, y) gode di una certa proprietà P .* Possiamo chiederci: A è *vero* o *falso*? Per rispondere a tale domanda è indispensabile chiarire una serie di circostanze: innanzitutto l’*“ambiente”* nel quale la formula A deve essere interpretata, cioè che tipo di individui viene rappresentato da x e da y ? Ad esempio x e y possono essere numeri naturali, o interi (o razionali, o reali, o complessi ...), o altre cose ancora. Si tratta

poi di dare un significato a $P(-, -)$: avendo scelto un universo numerico al primo passo, potremmo scegliere per $P(-, -)$ la relazione $x + y = 0$ (ma potremmo scegliere tutt'altra cosa). A questo punto la formula atomica $P(x, y)$ sarà vera per una certa interpretazione e falsa in altre. Ad esempio, è vera se consideriamo i numeri interi, la proprietà $x + y = 0$ e l'assegnazione del valore 3 ad x e del valore -3 ad y ; ma non è vera se, lasciando invariate le prime due scelte, assegniamo 3 ad x e 4 ad y . La situazione è ancora diversa se cambiamo la scelta dell'universo e consideriamo i numeri naturali. In quel caso nessuna assegnazione rende vera la formula atomica $P(x, y)$.

Notiamo che, nella formula $\forall x \exists y P(x, y)$, il valore di verità varia in base all'universo ed all'interpretazione del predicato, e non più rispetto

all’assegnazione di valori alle variabili, perchè queste sono ormai “vincolate” dai quantificatori. Con la fissata interpretazione di $P(-, -)$, la formula è falsa nei naturali e vera in tutti gli altri dominii numerici considerati.

Come nella logica degli enunciati, ci saranno poi formule vere indipendentemente dall’interpretazione specifica, ma soltanto a causa della loro struttura sintattica e delle regole generali che daremo per interpretarle. Vediamo qui alcuni importanti esempi, che ci serviranno per fare altre considerazioni. Dire che esiste almeno un x per il quale è verificata la proprietà P equivale a dire che non per ogni x per la proprietà P risulta non verificata. In simboli:

$$\exists x P \leftrightarrow (\neg \forall x \neg P)$$

Ciò vuol dire che il quantificatore

esistenziale \exists può essere definito a partire dal quantificatore \forall e dall'operatore di negazione. Osserviamo che, analogamente, il quantificatore universale \forall potrebbe essere definito a partire dal quantificatore \exists e dall'operatore di negazione: dire che per ogni x è verificata la proprietà P equivale infatti a dire che non esiste alcun x per il quale la proprietà P risulta non verificata, in simboli

$$\forall x P \leftrightarrow (\neg \exists x \neg P)$$

Ad esempio, al posto degli enunciati

*Esiste (almeno) un quadrilatero equilatero
Ogni italiano è europeo*

possiamo scrivere

Non ogni quadrilatero è non equilatero
Non esiste alcun italiano che non sia europeo

Ciò consente di precisare alcune importanti osservazioni riguardanti la negazione di una frase quantificata, che riassumiamo così (ricordiamo che $\neg\neg A$ equivale ad A):

- la negazione di $\exists xP$ è $\neg(\neg\forall x \neg P)$, ovvero è $\forall x \neg P$;
- la negazione di $\forall xP$ è $\neg(\neg\exists x \neg P)$, ovvero è $\exists x \neg P$.

Esempio 9.1. *Le negazioni dei due enunciati:*

Esiste (almeno) un quadrilatero equilatero
Ogni italiano è europeo

sono, rispettivamente, gli enunciati:

*Ogni quadrilatero non è equilatero
Esiste (almeno) un italiano che non è
europeo*

Nota 9.1. *Le frasi sopra scritte sono state ottenute combinando alcuni termini che abbiamo introdotto in ambito logico; ma in italiano, ad esempio, l'espressione “ogni quadrilatero non è equilatero” non viene usata. Essa è sempre sostituita dalla frase (con lo stesso significato) “nessun quadrilatero è equilatero”.*

9.1.3 Variabili vincolate, variabili libere e sostituzioni

Abbiamo notato nel precedente paragrafo che è molto importante, ai fini della

determinazione della verità di una formula, sapere se questa dipenda o meno dall’assegnazione di valori alle variabili e che questo accade a seconda se la variabile è vincolata o meno da un quantificatore. Dobbiamo ora precisare questo concetto.

Definizione 9.1. *Nella formula $\forall x \alpha$, α si chiama campo d’azione o ambito del quantificatore \forall ed analogamente diremo per \exists . Nelle formule $\forall x \alpha$ e $\exists x \alpha$ la variabile x si dice vincolata (o quantificata o legata).*

Specifichiamo che un’occorrenza di una variabile x si dice vincolata quando essa è la variabile che segue un \forall oppure è nel campo d’azione di un tale quantificatore; lo stesso vale per \exists . Un’occorrenza di una variabile non vincolata si dice *libera*. Si noti dunque che all’interno di una formula possono esserci sia occorrenze libere che

occorrenze vincolate di x . Tale è ad esempio la x nella formula $\alpha(x) \vee \forall x\beta(x)$: la prima occorrenza della variabile x è libera, la seconda vincolata.

Non sempre espliciteremo tutte le variabili che compaiono in una formula. Tuttavia spesso se una variabile z appare libera nella formula α si scrive $\alpha(z)$. L'insieme delle variabili nel termine t si indica con $\text{var}(t)$, mentre l'insieme delle variabili libere della formula α si indica con $\text{free}(\alpha)$. Diamo anche delle variabili libere una definizione induttiva. Siano '*' un connettivo binario e Q un quantificatore; allora

1. le variabili libere di α sono $\text{free}(\alpha)$;
2. le variabili libere di $\alpha * \beta$ sono quelle di α unite a quelle di β , cioè $\text{free}(\alpha * \beta) = \text{free}(\alpha) \cup \text{free}(\beta)$;
3. le variabili libere di $Qx \alpha$ sono quelle di

α tranne x , cioè $\text{fr ee}(Qx \alpha) = \text{fr ee}(\alpha) \setminus \{x\}$.

Definizione 9.2. *Un termine senza variabili si dice chiuso. Una formula predicativa si dice chiusa se non ha variabili libere, aperta altrimenti.*

La definizione di variabile vincolata è ispirata alle notazioni con variabile *apparente* in matematica: la variabile x in $\lim_{x \rightarrow c} f(x)$, la variabile n in $\lim_{n \rightarrow +\infty} a_n$, la variabile x in $\int_a^b f(x)dx$, etc. Osserviamo che l'espressione “variabile apparente” (originariamente contrapposta a “variabile reale”) fu introdotta nel 1897 da Giuseppe Peano che scrive:

“In queste spiegazioni diciamo che una lettera che compare in una formula è *reale* oppure *apparente*, a seconda che il valore della formula dipenda o non dipenda dal

nome di questa lettera. Così, in $\int_0^1 x^m dx$ la lettera x è apparente e la lettera m reale. Tutte le lettere che compaiono in un teorema sono apparenti, perchè la sua verità è indipendente dai nomi delle lettere.”

(da *Formulaire*, 2, 1, 23)

Il significato di una formula non varia al variare delle variabili vincolate e, quindi, non ne dipende. Perciò la sostituzione di una variabile vincolata con un’altra che non compaia già nella formula, non cambia certo il significato della formula. Infatti, la formula $\forall x P(x)$ sarà perfettamente equivalente a $\forall y P(y)$, purchè y non comparisse già nell’espressione: in quel caso si verrebbe a creare un vincolo che non esisteva prima.

La questione della possibilità di cambiare il nome di una variabile libera o,

più generalmente, di sostituirla con un altro termine, è abbastanza delicata.

Definizione 9.3. *Il termine t si dice liberamente sostituibile a x nella formula $a(x)$ se sostituendo t in tutte le occorrenze libere di x in $a(x)$ nessuna variabile di t risulta quantificata dopo la sostituzione. In questo caso si scrive $a(t)$.*

Naturalmente la sostituzione di un termine al posto di una variabile libera in una formula cambia il significato della formula stessa perché ne restringe la variabilità. Intuitivamente, se una formula sarà vera in presenza di una variabile libera, sarà ancora vera quando questa verrà sostituita da un termine, perché avremo a che fare con un’istanza particolare della formula, mentre il contrario non sarà garantito. Tuttavia, per alcuni scopi, ad esempio per la

soddisfabilità, potremo ricorrere a particolari sostituzioni.

Esempio 9.2. (*Controesempio*).

Scrivendo

$$\exists y(\neg(x = y))$$

diciamo che esiste un y diverso da x : affermazione plausibile, ad esempio in un insieme numerico costituito da più elementi distinti. Attenzione: il termine y non è liberamente sostituibile a x . Se sostituissimo y al posto di x otterremmo infatti

$$\exists y(\neg(y = y))$$

che è evidentemente sempre falsa.

9.1.4 Modelli e validità

Abbiamo visto che la verità di una formula dipende dall'universo o *dominio* D nel

quale viene fatta un'assegnazione di valore alle variabili libere e dall'interpretazione che diamo ai simboli predicativi e funzionali che compaiono nella formula.

Definizione 9.4. *Data una formula F , un'interpretazione M per F consiste in:*

1. *un dominio non vuoto D ;*
2. *una relazione n -aria $|P| \subseteq D^n$ per ogni simbolo predicativo n -ario P che compare in F ;*
3. *una funzione n -aria $|f| \subseteq D^n \rightarrow D$ per ogni simbolo funzionale n -ario f che compare in F ;*
4. *un elemento fissato di D per ogni simbolo di costante che compare in F .*

Le variabili sono da pensarsi interpretate genericamente da elementi di D ; i temini saranno quindi induttivamente

interpretati come elementi di D , più o meno determinati a seconda di quante variabili libere contengano.

Esempio 9.3. Consideriamo la formula $\forall x P(x, f(x))$. Prendiamo come dominio di interpretazione l'insieme dei numeri naturali, come interpretazione del predicato binario P la relazione \leq e come interpretazione del simbolo funzionale f la funzione successore; alla variabile x si assegnerà un numero generico n . La formula verrà allora interpretata come segue: “per ogni $n \in N$, $n \leq n + 1$ ”. Con questa interpretazione la formula è vera, ma potremmo cambiare qualcosa e farla diventare falsa: basterebbe interpretare P con \geq , mantenendo tutto il resto inalterato.

Supponiamo invece, sempre nel dominio dei naturali, di interpretare P con la relazione di divisibilità. La formula

risulta di nuovo falsa ma, se consideriamo soltanto la sottoformula aperta $P(x, f(x))$, questa sarà talvolta vera, talaltra falsa, a seconda del valore che assegniamo alla variabile x . Si provi ad assegnare prima il valore 1 e poi il valore 2.

Formalizziamo ora quanto osservato nell'esempio precedente.

Definizione 9.5. *Data un'interpretazione M per F , F si dice soddisfacibile in M se esiste un'assegnazione di valori alle variabili libere in F tale che:*

1. *se F è atomica, cioè del tipo $P(t_1, \dots, t_n)$, allora deve essere $(|t_1|, \dots, |t_n|) \in |P|$;*
2. *se F è del tipo $H \wedge G$, allora devono essere soddisfatte H e G ;*
3. *se F è del tipo $H \vee G$, allora deve*

- essere soddisfatta o H o G ;*
4. *se F è del tipo $H \rightarrow G$, allora se è soddisfatta H allora deve esserlo anche G ;*
 5. *se F è del tipo $\neg H$, allora non deve essere soddisfatta H ;*
 6. *se F è del tipo $\forall xH(x)$, allora ogni assegnazione di valore alla x deve soddisfare H ;*
 7. *se F è del tipo $\exists xH(x)$, allora esiste un'assegnazione di valore alla x che soddisfa H .*

Data un'interpretazione M per F , F si dice vera in M se è soddisfatta per ogni assegnazione di valore alle variabili libere; M si dice allora modello di F , in simboli $M \models F$.

Una formula F , si dice (logicamente) valida se risulta vera in ogni interpretazione. Si scrive allora $\models F$.

Analogamente, sia Φ un insieme di formule chiuse (che talvolta viene detto *teoria*: v. più avanti); la scrittura

$$M \models \Phi$$

esprime la verità di tutte le formule chiuse di Φ in M . Diremo inoltre che F è *conseguenza logica* di Φ quando F vale in tutti i modelli in cui valgono tutte le formule di Φ . Ciò si esprime scrivendo

$$\Phi \models F$$

Infine, se con il simbolo \perp indichiamo un assurdo, il fatto che Φ non sia soddisfacibile si esprime scrivendo:

$$\Phi \models \perp$$

Esempio 9.4. *Esprimiamo attraverso la simbologia introdotta il principio della dimostrazione per assurdo nella sua forma generale:*

$$\Phi \models F \text{ se e solo se } \Phi \cup \{\neg F\} \models \perp$$

Proposizione 9.1. *Le formule seguenti sono valide nel calcolo dei predicati (esercizio):*

$$\forall x P(x) \leftrightarrow \neg \exists x \neg P(x)$$

$$\exists x P(x) \leftrightarrow \neg \forall x \neg P(x)$$

$$\forall x P(x) \rightarrow \exists x P(x)$$

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

$$\forall x \forall y P(x, y) \leftrightarrow \forall y \forall x P(x, y)$$

$$\exists x \exists y P(x, y) \leftrightarrow \exists y \exists x P(x, y)$$

$$\exists x (P(x) \vee Q(x)) \leftrightarrow (\exists x P(x) \vee \exists x Q(x))$$

$$\forall x (P(x) \wedge Q(x)) \leftrightarrow (\forall x P(x) \wedge \forall x Q(x))$$

$$(\forall x P(x) \vee \forall x Q(x)) \rightarrow \forall x (P(x) \vee Q(x))$$

$$\exists x (P(x) \wedge Q(x)) \rightarrow (\exists x P(x) \wedge \exists x Q(x))$$

$$\forall x (P(x) \leftrightarrow Q(x)) \rightarrow (\forall x P(x) \leftrightarrow \forall x Q(x))$$

$Q(x))$

$\forall x (P(x) \leftrightarrow Q(x)) \rightarrow (\exists x P(x) \leftrightarrow \exists x Q(x))$

$(\exists x P(x) \vee Q) \leftrightarrow \exists x (P(x) \vee Q)$

$(\forall x P(x) \vee Q) \leftrightarrow \forall x (P(x) \vee Q)$

$(\exists x P(x) \wedge Q) \leftrightarrow \exists x (P(x) \wedge Q)$

$(\forall x P(x) \wedge Q) \leftrightarrow \forall x (P(x) \wedge Q)$

$\forall x (P \rightarrow Q(x)) \leftrightarrow (P \rightarrow \forall x Q(x))$

$\forall x (P(x) \rightarrow Q) \leftrightarrow (\exists x P(x) \rightarrow Q)$

$\exists x (P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \exists x Q(x))$

$(\exists x P(x) \rightarrow \forall x Q(x)) \rightarrow \forall x (P(x) \rightarrow Q(x))$

$\forall x (P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))$

$\forall x (P(x) \rightarrow Q(x)) \rightarrow (\exists x P(x) \rightarrow \exists x Q(x))$

$\forall x (P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \exists x Q(x))$

9.2 Il metodo dei tableau per il calcolo dei predicati

9.2.1 Tableau e quantificatori

La logica dei predicati si presenta come un completamento della logica degli enunciati; dunque costruiremo i tableau di formule predicative applicando innanzitutto le regole precedentemente introdotte per i tableau proposizionali.

Do-vremo però introdurre altre regole per poter trattare anche i quantificatori; per tale ragione, questa parte della logica può essere anche chiamata *calcolo dei quantificatori*.

Introduciamo la questione occupandoci dell'esempio seguente, nel quale cercheremo di costruire il tableau della negazione della formula predicativa valida $\forall x (P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x$

$Q(x))$:

$$\begin{array}{c} \neg(\forall x(P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))) \\ | \\ \forall x(P(x) \rightarrow Q(x)), \neg(\forall x P(x) \rightarrow \forall x Q(x)) \\ | \\ \forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg\forall x Q(x) \end{array}$$

Il nostro lavoro non è evidentemente finito: il tableau così ottenuto non è chiuso e non abbiamo eliminato i quantificatori. Esaminiamo l'ultima formula scritta

$$\begin{array}{c} \neg(\forall x(P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))) \\ | \\ \forall x(P(x) \rightarrow Q(x)), \neg(\forall x P(x) \rightarrow \forall x Q(x)) \\ | \\ \forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg\forall x Q(x) \\ | \\ \forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg Q(a) \end{array}$$

che sappiamo essere equivalente a

$$\exists x \neg Q(x)$$

In fondo la negazione di una formula quantificata universalmente è una formula quantificata esistenzialmente,

La costruzione di un tableau, come

sappiamo, formalizza la ricerca di un controesempio; dunque a questo punto dobbiamo sostituire alla variabile x un qualche elemento del dominio nel quale viene considerata la formula, tentando di arrivare ad una contraddizione; indichiamo con a tale elemento (*istanza*):

$$\begin{array}{c}
 \neg(\forall x(P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))) \\
 | \\
 \forall x(P(x) \rightarrow Q(x)), \neg(\forall x P(x) \rightarrow \forall x Q(x)) \\
 | \\
 \forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg\forall x Q(x) \\
 | \\
 \forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg Q(a)
 \end{array}$$

Per quanto riguarda i quantificatori universali, possiamo scrivere le due formule $\forall x (P(x) \rightarrow Q(x))$ e $\forall x P(x)$ facendo proprio riferimento all'elemento a considerato (esse, in quanto universali, sono riferite a *tutti* gli elementi!), dunque scrivendo $P(a) \rightarrow Q(a)$ e $P(a)$. Otteniamo quindi

$$\begin{array}{c}
\neg(\forall x(P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))) \\
| \\
\forall x(P(x) \rightarrow Q(x)), \neg(\forall x P(x) \rightarrow \forall x Q(x)) \\
| \\
\forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg\forall x Q(x) \\
| \\
\forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg Q(a) \\
| \\
\neg Q(a), P(a) \rightarrow Q(a), P(a)
\end{array}$$

Anticipiamo qui che questo modo di procedere è piuttosto ingenuo e che, in generale, sarà necessario, su questa procedura, fare una precisazione che, nel caso ora in esame, non risulta indispensabile.

L'applicazione della regola proposizionale per $P(a) \rightarrow Q(a)$ porta a

$$\begin{array}{c}
\neg(\forall x(P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))) \\
| \\
\forall x(P(x) \rightarrow Q(x)), \neg(\forall x P(x) \rightarrow \forall x Q(x)) \\
| \\
\forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg\forall x Q(x) \\
| \\
\forall x(P(x) \rightarrow Q(x)), \forall x P(x), \neg Q(a) \\
| \\
\neg Q(a), P(a) \rightarrow Q(a), P(a) \\
/ \quad \backslash \\
\neg Q(a), \neg P(a), P(a) \quad \neg Q(a), Q(a), P(a) \\
\Diamond \qquad \qquad \qquad \Diamond
\end{array}$$

Il tableau predicativo così ottenuto è chiuso.

Come accennato, il procedimento sopra esemplificato necessita però di alcune precisazioni fondamentali. Anticipiamo le due più importanti:

1. la scelta di un elemento particolare nel caso di una formula con un quantificatore universale deve poter essere ripetuta più volte, dal momento che ogni elemento può essere utilizzato;
2. la scelta di un elemento particolare nel caso di formule con quantificatori esistenziali deve essere indipendente da elementi precedentemente scelti, perché l'esistenza di un elemento che verifica una certa proprietà non implica che un elemento già dato la verifichi.

Il seguente esempio renderà chiara l'ultima precisazione sulle regole

predicative per la costruzione dei tableau.

Esempio 9.5 (Controesempio).

Consideriamo le ipotesi del Teorema di Cauchy:

$$\left\{ \begin{array}{l} f \text{ continua in } [a, b] \\ f \text{ derivabile in }]a, b[\\ g \text{ continua in } [a, b] \\ g \text{ derivabile in }]a, b[\\ \forall x \in]a, b[.g'(x) \neq 0 \end{array} \right.$$

In tale situazione, le ipotesi del Teorema di Lagrange sono rispettate sia dalla funzione f che dalla g ; potremmo quindi scrivere

$$\exists c \in]a, b[.f'(c) = \frac{f(b) - f(a)}{b - a}$$

$$\exists c \in]a, b[.g'(c) = \frac{g(b) - g(a)}{b - a}$$

$$\exists c \in]a, b[.\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

e, dividendo membro a membro, finiremmo con l'ottenere

$$\exists c \in]a, b[. f'(c) = \frac{f(b) - f(a)}{b - a}$$

$$\exists d \in]a, b[. g'(d) = \frac{g(b) - g(a)}{b - a}$$

cioè la tesi del Teorema di Cauchy! La “dimostrazione” riportata non è però accettabile: la scrittura corretta di quanto ottenuto applicando due distinte volte il Teorema di Lagrange dovrebbe essere, per quanto riguarda l’azione dei due quantificatori esistenziali:

$$\begin{array}{ccc} \forall x A(x) & & \neg \exists x B(x) \\ | & & | \\ \forall x A(x), A(a) & & \neg \exists x B(x), \neg B(a) \end{array}$$

(con c e d non necessariamente uguali) e da ciò non può direttamente essere ottenuta la tesi del Teorema di Cauchy (che afferma l’esistenza “di un (singolo) punto c dell’intervallo]a, b[tale che...”).

L’esempio che sarà proposto nel paragrafo seguente riprenderà la situazione ora rilevata nel metodo dei

tableau.

9.2.2 Regole per la costruzione di un tableau predicativo

Come sopra osservato, tutte le regole introdotte precedentemente con riferimento ai tableau proposizionali possono essere nuovamente applicate per la costruzione di tableau predicativi. Ad esse vanno aggiunte due nuove regole:

1. (γ -regole) se una formula di un ramo di un tableau è del tipo $\forall x A(x)$ oppure $\neg \exists x B(x)$, aggiungiamo un nuovo nodo:

$$\begin{array}{ccc} \forall x A(x) & & \neg \exists x B(x) \\ | & & | \\ \forall x A(x), A(a) & & \neg \exists x B(x), \neg B(a) \end{array}$$

2. (δ -regole) se una formula di un ramo di un tableau è del tipo $\exists x A(x)$ oppure $\neg \forall x B(x)$, aggiungiamo un nuovo nodo:

$$\begin{array}{ccc} \exists x A(x) & & \neg \forall x B(x) \\ | & & | \\ A(a) & & \neg B(a) \end{array}$$

purchè la costante a non sia mai stata usata nei nodi precedenti.

Diversamente dalle δ -regole e da quelle relative ai tableau proposizionali, le γ regole pongono sì una formula più semplice sul nuovo nodo, ma mantengono anche la formula quantificata (nella costruzione illustrata nel paragrafo precedente non abbiamo fatto ciò, ma in tale caso, come sopra segnalato, questa omissione non è stata tale da pregiudicare la chiusura del tableau). Ciò accade per garantire l'essenziale possibilità di ripetere l'istanziazione della formula quantificata (universalmente) anche a nuove costanti che vengano introdotte successivamente.

Questa osservazione ha un'importante

conseguenza: mentre nell’ambito della logica degli enunciati il metodo dei tableau è di tipo *decisionale* (il tableau risulta comunque finito, chiuso o non chiuso e quindi fornisce una risposta al fatto che la negazione della formula esaminata sia o non sia una tautologia in un numero finito di passi), in ambito predicativo tale metodo può portare a tableau infiniti (e quindi può *non* portare a tale risposta in un numero finito di passi: riprenderemo questa considerazione nella Sezione 9.2.3).

Inoltre, ripetiamo ancora quanto anticipato a proposito della presenza di formule quantificate esistenzialmente: la scelta di un elemento particolare in tale caso deve *essere del tutto indipendente da altre istanziazioni*. Illustriamo questa osservazione con il seguente esempio.

Esempio 9.6 (Controesempio).

Consideriamo il seguente tableau, riferito ad una formula con due quantificatori esistenziali:

$$\begin{array}{c} \exists x P(x) \wedge \exists x \neg P(x) \\ | \\ \exists x P(x), \exists x \neg P(x) \\ | \\ P(a), \neg P(a) \end{array}$$

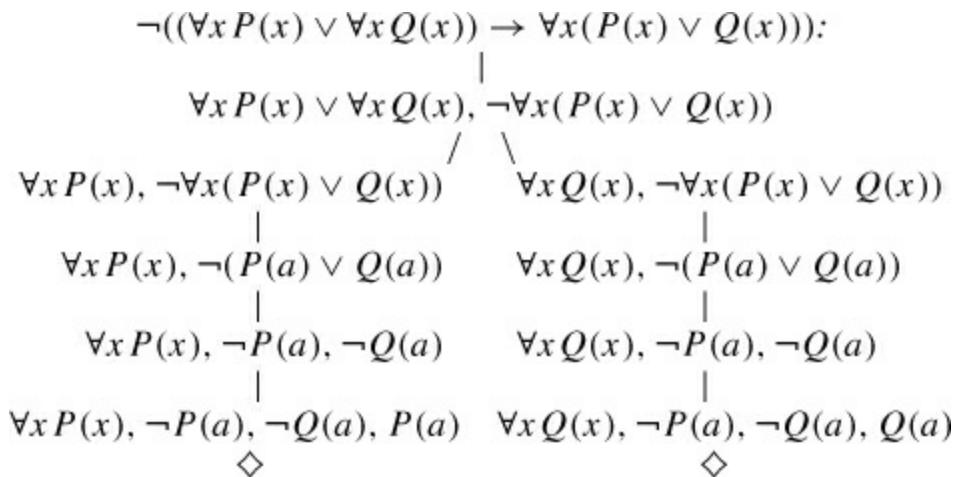
Potremmo essere tentati di affermare la chiusura di tale tableau, ma sarebbe una conclusione errata: l'esistenza di un x per cui è $P(x)$ e di un x per cui è $\neg P(x)$ non implica che tali x siano lo stesso elemento (a meno che ad x non sia imposto di variare in un dominio costituito da un solo elemento). Pertanto l'ultimo nodo avrebbe dovuto essere:

$$P(a), \neg P(b)$$

e in questo caso il tableau non può essere considerato chiuso.

Le formule elencate come valide nel calcolo dei predicati alla fine del precedente capitolo possono essere verificate, per esercizio, con il metodo dei tableau predicativi. Ne diamo qui due esempi.

Esempio 9.7. *Costruiamo il tableau predicativo per la negazione della formula $(\forall x P(x) \vee \forall x Q(x)) \rightarrow \forall x (P(x) \vee Q(x))$:*



Il tableau predicativo così ottenuto è chiuso; dunque, la formula $(\forall x P(x) \vee \forall x Q(x)) \rightarrow \forall x (P(x) \vee Q(x))$ è valida. Si

noti che abbiamo sostituito la stessa a in due formule esistenziali, ma queste si trovavano su rami diversi, quindi su due diverse “ricerche di controesempio” (corrispondenti a due diverse interpretazioni).

Esempio 9.8. Costruiamo il tableau predicativo per la negazione della formula $\forall x (P(x) \leftrightarrow Q(x)) \rightarrow (\exists x P(x) \leftrightarrow \exists x Q(x))$:

| | |
|---|---|
| $\neg(\forall x (P(x) \leftrightarrow Q(x)) \rightarrow (\exists x P(x) \leftrightarrow \exists x Q(x)))$: | |
| | |
| $\forall x (P(x) \leftrightarrow Q(x)), \neg(\exists x P(x) \leftrightarrow \exists x Q(x))$ | |
| | / \ |
| $\forall x (P(x) \leftrightarrow Q(x)),$ $\exists x P(x), \neg\exists x Q(x)$ | $\forall x (P(x) \leftrightarrow Q(x)),$ $\neg\exists x P(x), \exists x Q(x)$ |
| | |
| $\forall x (P(x) \leftrightarrow Q(x)),$ $P(a), \neg\exists x Q(x)$ | $\forall x (P(x) \leftrightarrow Q(x)),$ $\neg\exists x P(x), Q(a)$ |
| | |
| $\forall x (P(x) \leftrightarrow Q(x)),$ $P(a), \neg\exists x Q(x), \neg Q(a)$ | $\forall x (P(x) \leftrightarrow Q(x)),$ $\neg\exists x P(x), Q(a), \neg P(a)$ |
| | |
| $\forall x (P(x) \leftrightarrow Q(x)),$ $P(a) \leftrightarrow Q(a), P(a),$ $\neg\exists x Q(x), \neg Q(a)$ | $\forall x (P(x) \leftrightarrow Q(x)),$ $P(a) \leftrightarrow Q(a), Q(a),$ $\neg\exists x P(x), \neg P(a)$ |
| | / \ / \ |
| $\forall x (P(x) \leftrightarrow Q(x)),$ $P(a), Q(a), P(a),$ $\neg P(a), \neg Q(a), P(a),$ $\neg\exists x Q(x), \neg Q(a)$ | $\forall x (P(x) \leftrightarrow Q(x)),$ $P(a), Q(a), Q(a),$ $\neg P(a), \neg Q(a), Q(a),$ $\neg\exists x P(x), \neg P(a)$ |
| \diamond | \diamond |

Il tableau predattivo così ottenuto è chiuso. Ciò ci consente di concludere che la formula $\forall(x P(x) \leftrightarrow Q(x)) \rightarrow (\exists x P(x) \leftrightarrow \exists x Q(x))$ è valida.

9.2.3 Correttezza e completezza del metodo dei tableau

Sia F una formula predattiva; usiamo la notazione $\vdash_T F$ per esprimere che F ha un tableau chiuso. Ricordiamo inoltre che la notazione $\models F$ esprime la validità di F . Usando tali notazioni, possiamo enunciare e dimostrare la correttezza e completezza del metodo dei tableau per il calcolo dei predicati.

Teorema 9.1 (Correttezza del metodo dei tableau). *Se $\vdash_T F$ allora $\models F$.*

Dimostrazione. La dimostrazione è per induzione sulla complessità della formula

e ricalca quella nel calcolo degli enunciati. Dovrà però essere estesa al caso nel quale si siano applicate le regole γ e δ .

Supponiamo allora che per un insieme di formule si abbia un tableau chiuso (di conseguenza finito).

Supponiamo di aver applicato una regola γ ad una formula che si trova sul nodo n di altezza h . Si ha allora che $U(n) = \{\forall x P(x)\} \cup U_0$ e $U(n') = \{\forall x P(x), P(a)\} \cup U_0$ (dove, al solito, U_0 può anche essere vuoto) e l'altezza di n è $h - 1$.

Dunque, per l'ipotesi induttiva, $U(n)$ è insoddisfacibile. Ciò significa che, data una qualsiasi interpretazione, o risulta falsa una formula di U_0 o $\forall x P(x)$ (e, quindi, anche $U(n)$ è insoddisfacibile), oppure risulta falsa $P(a)$; ma allora anche $\forall x P(x)$ risulterebbe falsa (da cui l'insoddisfacibilità di $U(n)$).

Se invece abbiamo applicato una regola

δ ad una formula che si trova sul nodo n di altezza h , allora $U(n) = \{ \exists x P(x) \} \cup U_0$ e $U(n') = \{ P(a) \} \cup U_0$ e l'altezza di n è $h - 1$. Dunque, per l'ipotesi induttiva, $U(n')$ è insoddisfacibile. Ciò significa che, data una qualsiasi interpretazione, o risulta falsa una formula di U_0 (e quindi $U(n)$ è insoddisfacibile), oppure risulta falsa $P(a)$ e, quindi, anche $\exists x P(x)$ (da cui l'insoddisfacibilità di $U(n)$): infatti, a è stata scelta in modo del tutto indipendente da altri valori e, perciò, la falsità della formula si darà indipendentemente dall'istanziazione che avremo fatto per il quantificatore esistenziale.

□

Teorema 9.2 (Completezza del metodo dei tableau). Se $\models F$ allora $\vdash_T F$. La dimostrazione in questo caso è

piuttosto complessa e viene tralasciata. Infatti dal momento che il tableau potrebbe continuare indefinitamente, per assicurarci la completezza del metodo, dovremmo accertarci che tutti i casi possibili siano stati prima o poi esaminati. Per questo avremmo bisogno della *costruzione sistematica* del tableau, cioè di un procedimento che assicuri un'applicazione sistematica delle regole. Non riteniamo opportuno, date le dimensioni e lo scopo di questo volume, illustrare dettagliatamente le caratteristiche di tale procedimento di costruzione; per questo rimandiamo a [6].

È tuttavia interessante osservare che, dal momento che una formula soddisfacibile ha sicuramente un tableau aperto, un ramo aperto in esso può aiutarci ad individuare un modello come nel caso proposizionale. Infatti,

considerando come dominio l'insieme delle costanti che intervengono nel ramo ed interpretando i predicati come relazioni sulle costanti, più o meno soddisfatte a seconda dei letterali che troviamo sul ramo, si può spesso facilmente costruire un modello.

Esempio 9.9. Si consideri la formula $\forall y \exists x P(x, y) \rightarrow \exists x \forall y P(x, y)$. Il suo tableau è il seguente:

$$\begin{array}{ccc}
 \forall y \exists x P(x, y) \rightarrow \exists x \forall y P(x, y) & & \\
 / \qquad \backslash & & \\
 \neg \forall y \exists x P(x, y) & & \exists x \forall y P(x, y) \\
 | & & | \\
 \neg \exists x P(x, b) & & \forall y P(a, y) \\
 | & & | \\
 \neg \exists x P(x, b), \neg P(a, b) & & \forall y P(a, y), P(a, b)
 \end{array}$$

Chiaramente si potrebbero creare nuove istanze delle formule universali utilizzando y -regole, ma, altrettanto chiaramente, nessuno dei due rami si chiuderà, dal momento che su uno dei due il predicato P risulterà sempre negato e

sull'altro mai. La formula $\forall y \exists x P(x, y) \rightarrow \exists x \forall y P(x, y)$ è pertanto soddisfacibile. Il ramo di destra ci suggerisce, ad esempio, il modello $D = \{a, b\}$ con $|P| = \{(a, b), (a, a)\}$, mentre quello di sinistra il modello $D = \{a, b\}$ con $|P| = \emptyset$.

9.3 Il sistema di Gentzen per il calcolo dei predicati

La deduzione nel sistema di Gentzen, che nel capitolo precedente era stata introdotta con riferimento ai soli enunciati, può essere estesa al calcolo dei predicati. Alle regole proposizionali si aggiungono le nuove regole per i quantificatori esistenziali e universali (quelle della prima riga possono essere denominate γ -regole, quelle della seconda δ -regole):

$$\frac{U \cup \{\exists x A(x), A(a)\}}{U \cup \{\exists x A(x)\}}$$

$$\frac{U \cup \{\neg \forall x A(x), \neg A(a)\}}{U \cup \{\neg \forall x A(x)\}}$$

$$\frac{U \cup \{A(a)\}}{U \cup \{\forall x A(x)\}}$$

$$\frac{U \cup \{\neg A(a)\}}{U \cup \{\neg \exists x A(x)\}}$$

Inoltre, similmente alla costruzione dei tableau, le δ -regole possono essere applicate solamente a condizione che la costante a non appaia in U ; ciò è richiesto per non imporre restrizioni sull'interpretazione della costante a .

Illustriamo l'applicazione del metodo mediante l'esempio seguente.

Esempio 9.10. *Deduciamo la validità della formula predicativa $\exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$.*

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

$$\begin{array}{c}
\neg \forall y P(a, y), \neg P(a, b), \neg P(a, a), \exists x P(x, b), P(a, b) \\
\hline
\neg \forall y P(a, y), \neg P(a, a), \exists x P(x, b), P(a, b) \\
\hline
\neg \forall y P(a, y), \exists x P(x, b), P(a, b) \\
\hline
\neg \forall y P(a, y), \exists x P(x, b) \\
\hline
\neg \forall y P(a, y), \forall y \exists x P(x, y) \\
\hline
\neg \exists x \forall y P(x, y), \forall y \exists x P(x, y) \\
\hline
\exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)
\end{array}$$

Similmente al caso del calcolo degli enunciati, esiste una dimostrazione di Gentzen per la disgiunzione metalinguistica delle formule in U (scritto $\vdash G U$) se e soltanto se è possibile ottenere un tableau semantico chiuso per la congiunzione metalinguistica dei complementi delle formule di U . Ciò significa che il sistema di Gentzen è corretto e completo (la dimostrazione si ottiene dualizzando quella per i tableau: esercizio).

Teorema 9.3 (Correttezza e completezza del metodo deduttivo di

Gentzen). $\vdash G U$ se e solo se $\models U$.

9.4 Cenni sul sistema di Hilbert per il calcolo dei predicati

Anche il sistema di Hilbert, che nel capitolo precedente era stato introdotto con riferimento ai soli enunciati, può essere esteso al calcolo dei predicati. Ci limiteremo a sviluppare la questione facendo riferimento al quantificatore universale; per quanto riguarda il quantificatore esistenziale è infatti sufficiente ricordare che

$$\frac{\vdash A(a)}{\vdash \forall x A(x)}$$

Dobbiamo quindi aggiungere due schemi di assiomi ed una nuova regola per trattare formule quantificate. I nuovi

assiomi (o schemi di assiomi) sono:

$$(\text{Assioma 4}) \vdash \forall x A(x) \rightarrow A(a)$$

$$(\text{Assioma 5}) \vdash \forall x(A \rightarrow B(x)) \rightarrow (A \rightarrow \forall x B(x))$$

dove il simbolo “ \vdash ” denota, come al solito, la dimostrabilità. La nuova regola di inferenza è detta di *generalizzazione*:

$$\frac{\vdash A(a)}{\vdash \forall x A(x)}$$

Attenzione! La regola di generalizzazione non è una sorta di inversa per l’assioma 4. Se così fosse, il ruolo del quantificatore universale verrebbe totalmente annullato: una proprietà varrebbe per tutti se e soltanto se valesse per un individuo; e ciò è manifestamente senza senso. Dov’è la differenza? Mentre nell’assioma 4 $A(a)$ rappresenta una semplice istanziazione della formula $A(x)$ (e, quindi, un caso particolare), nella premessa della regola di

generalizzazione si dice che $A(a)$ deve essere un teorema, quindi valido per una *generica* costante a . Non stupisce perciò che da questa premessa si possa dedurre la validità di $\forall x A(x)$. L'argomentazione verrà ripresa formalmente nel teorema di correttezza.

Vista la sua grande utilità nello sviluppare prove, vorremmo ora poter estendere al caso predicativo la regola di deduzione che valeva nel caso proposizionale: per un insieme di formule U ed una coppia di formule A e B , vorremmo che valesse

$$\frac{U \cup \{A\} \vdash B}{U \vdash A \rightarrow B}$$

Nel caso predicativo, però, essa non vale in generale. Infatti potrà essere utilizzata purchè nella dimostrazione di $U \cup \{A\} \rightarrow B$ non sia applicata la regola di generalizzazione ad una costante che compare

in A . Questa condizione è strettamente legata a quelle sulla genericità della costante che abbiamo visto nelle δ -regole per i tableau ed per il sistema di Gentzen. Si prova (*teorema di deduzione*) che la regola di deduzione, con questa restrizione, è una regola derivata corretta.

Esempio 9.11 (Controesempio).

Supponendo che la regola di deduzione valga in generale (cioè, senza la restrizione sulle costanti generalizzate), riusciamo a dimostrare l'implicazione inversa dell'Assioma 4, cioè $A(a) \rightarrow \forall x A(x)$ che, sappiamo già, renderebbe scorretto l'intero sistema deduttivo:

1. *per ipotesi, $\{A(a)\} \rightarrow A(a)$;*
2. *per generalizzazione applicata al punto 1, $\{A(a)\} \rightarrow \forall x A(x)$;*
3. *per deduzione (errata!), $\rightarrow A(a) \rightarrow \forall x A(x)$.*

L'errore consiste nell'aver usato la generalizzazione sulla costante a che compare nell'ipotesi $A(a)$.

Illustriamo invece un'applicazione corretta del metodo di Hilbert mediante l'esempio seguente.

Esempio 9.12. *Deduciamo la validità della formula $\forall x P(x) \rightarrow Q(x)) \rightarrow (\forall x p(x) \rightarrow \forall x (Q(x))$*

1. per ipotesi, $\{\forall x (P(x) \rightarrow Q(x)), \forall x P(x)\} \rightarrow \forall x P(x);$

2. per l'assioma 4 e modus ponens applicato con ipotesi 1,

$$\{\forall x (P(x) \rightarrow Q(x)), \forall x P(x)\} \vdash P(a)$$

3. per ipotesi, $\{\forall x (P(x) \rightarrow Q(x)), \forall x P(x)\} \rightarrow \forall x (P(x) \rightarrow Q(x));$

4. ancora per l'assioma 4 e modus ponens

applicato con ipotesi 3,

$$\{\forall x (P(x) \rightarrow Q(x)), \forall x P(x)\} \rightarrow P(a) \rightarrow Q(a)$$

5. *per modus ponens applicato a 2 e 4,*
 $\{\forall x (P(x) \rightarrow Q(x)), \forall x P(x)\} \rightarrow Q(a);$
6. *per generalizzazione di 5,* $\{\forall x (P(x) \rightarrow Q(x)), \forall x P(x)\} \rightarrow \forall x Q(x);$
7. *per deduzione,* $\{\forall x (P(x) \rightarrow Q(x))\} \rightarrow \forall x P(x) \rightarrow \forall x Q(x);$
8. *per deduzione,* $\rightarrow \forall x (P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x)).$

Si noti che, come nel caso proposizionale, quanto provato nell'esempio precedente può legittimamente trasformarsi in una nuova regola (derivata), spesso denominata ancora “generalizzazione”:

$$\frac{\vdash A(a) \rightarrow B(a)}{\vdash \forall x A(x) \rightarrow \forall x B(x)}$$

Di nuovo, si può dimostrare che A è valida se e soltanto se c'è una sua dimostrazione nel sistema di Hilbert (scritto $\rightarrow H A$); ciò significa che il sistema di Hilbert è corretto e completo, come esprimono i teoremi seguenti.

Teorema 9.4 (Correttezza del metodo deduttivo di Hilbert). *Se $\vdash H A$ allora $\models A$.*

Dimostrazione. Il teorema di correttezza può essere dimostrato estendendo la prova del caso proposizionale con l'uso le interpretazioni insiemistiche al posto delle tavole di verità. Dovremo far vedere che i due nuovi assiomi sono veri per ogni interpretazione. Consideriamo l'Assioma 4 (la verifica per l'Assioma 5 viene lasciata come esercizio): si tratta di

un'implicazione; quindi, qualunque sia l'interpretazione data, ogniqualvolta risulti vero l'antecedente, deve risultare vero il conseguente. Ma se l'interpretazione della formula A risulta vera per ogni elemento del dominio di interpretazione, risulterà in particolare vera per l'elemento associato nell'interpretazione alla costante a .

Per la regola di generalizzazione, come già osservato precedentemente, se $A(a)$ è un teorema e, quindi, vero (per ipotesi induttiva) in ogni interpretazione, esso sarà vero al variare dell'interpretazione di a in un qualunque fissato dominio; $A(x)$ sarà perciò vero qualunque sia l'assegnazione di valore alla variabile x in quel dominio. Sarà allora valido $\forall x A(x)$.

□

Teorema 9.5 (Completezza del metodo deduttivo di Hilbert). Se $\models A$

allora $\vdash_H A$.

Del teorema di completezza non diamo dimostrazione, dati i limiti di questo testo; chi è interessato può consultare [6] o un altro manuale di logica. Il risultato può essere facilmente generalizzato considerando un particolare sistema di ipotesi (teoria), ottenendo

$$\Phi \models A \text{ se e soltanto se } \Phi \vdash_H A$$

9.5 Teorie

Dato un linguaggio come quello che abbiamo visto per il calcolo dei predicati, una *teoria* è data da un insieme di assiomi che vengono espressi come formule chiuse di quel linguaggio, supponendo che quest'ultimo comprenda simboli predicativi, funzionali e costanti sufficienti ad esprimerli. Una volta

aggiunto un apparato dimostrativo opportuno, ad esempio il sistema deduttivo di Hilbert con il suo sistema di assiomi e regole, siamo pronti a dimostrare teoremi di quella teoria, cioè formule che possono essere derivate assumendo gli assiomi specifici.

Un primo esempio di teoria è dato dalla geometria euclidea con i suoi cinque, ben noti, assiomi. Qui potremmo pensare di presentare in maniera del tutto formale la teoria dei numeri naturali, come l'abbiamo introdotta nel [capitolo 5](#). Avremmo bisogno di un linguaggio che contenga un predicato binario “ $= (-, -)$ ” (più comunemente scritto “ $- = -$ ”) da interpretarsi come uguaglianza, simboli funzionali per il successore “ $s(-)$ ”, la somma “ $- + -$ ”, il prodotto “ $- * -$ ” e una costante “ o ” per lo zero. L'insieme degli assiomi assumerebbe la seguente forma,

dove ogni formula è da intendersi quantificata universalmente:

1. $x = x$
2. $x = y \rightarrow y = x$
3. $x = y \rightarrow (y = z \rightarrow x = z)$
4. $x = y \rightarrow s(x) = s(y)$
5. $s(x) = s(y) \rightarrow x = y$
6. $0 = s(x)$
7. $x + 0 = x$
8. $x + s(y) = s(x + y)$
9. $x * 0 = 0$
10. $x * s(y) = x * y + x$
11. Per ogni formula $A(x)$ con simboli della teoria

$$A(0) \rightarrow (\forall x(A(x) \rightarrow A(s(x))) \rightarrow \forall x A(x))$$

I primi tre assiomi impongono al predicato “=” di essere interpretato mediante una relazione di equivalenza,

come è giusto per un'uguaglianza; il quarto richiede che il successore sia veramente una funzione e con il quinto chiediamo che la funzione sia iniettiva. Il sesto ci dice che lo o non è successore. Gli altri sono la controparte della definizione per induzione delle operazioni di somma e prodotto. Un discorso particolare richiede il principio di induzione, che nel [capitolo 5](#) abbiamo presentato come assioma: esso non può essere presentato come una formula del calcolo dei predicati, perché comprende una quantificazione non relativa ad una variabile, ma ad una formula. Per ovviare a questo problema, lo abbiamo qui sostituito con un'intera famiglia di formule, ottenendo così un sistema infinito di assiomi per l'aritmetica di Peano. Da questo sistema possiamo pensare di dedurre come teoremi tutte le proprietà dei numeri.

In realtà le cose non stanno esattamente in questo modo. L'intento primario di Hilbert nel dare una rigorosa impostazione assiomatica alla logica ed alla geometria era stato quello di evitare a priori le antinomie che si erano prodotte in matematica alla fine del secolo XIX (v. il paragrafo 4.4). Ripercorriamo allora, per concludere questa parte, il ragionamento che portò Kurt Gödel [20] a spegnere ogni speranza in tale senso per tutte le teorie che contenessero l'aritmetica.

Osserviamo che il linguaggio di una teoria, come l'abbiamo appena descritta, contiene soltanto un numero finito di simboli e che le formule ben formate sono costituite ciascuna da un numero finito di essi; sarà dunque possibile assegnare ad ogni simbolo e ad ogni formula un numero naturale. Occorrerà, però, che l'assegnazione sia univoca nei due sensi,

cioè che, data una formula o un simbolo, si possa determinare in un sol modo il numero che gli compete e, viceversa, dato un numero, sia determinato (se esiste) l'elemento del linguaggio da esso contraddistinto. Per fare questo sfruttiamo la proprietà di fattorizzazione unica dei numeri naturali (v. Proposizione 5.3). Ai segni elementari costanti dell'alfabeto si fanno per esempio corrispondere i primi numeri naturali, diciamo fino a n , alle variabili i numeri primi più grandi di n , ai simboli di enunciato i quadrati di tali numeri, ai simboli di predicato i cubi. Ad una formula viene associato il prodotto di tanti numeri primi in successione rigorosa quanti sono i simboli che vi compaiono, ognuno di essi con esponente uguale al numero associato al simbolo (numero di Gödel). Questa codifica (o una equivalente) può essere estesa anche alle successioni di

formule: data una successione di k formule, ad essa associamo il prodotto di k numeri primi in successione rigorosa ognuno di essi con esponente uguale al numero associato alla corrispondente formula. Si può verificare facilmente che in questo modo non ogni numero naturale sarà un numero di Go[“] del, ma sarà sempre possibile determinare se siamo in tale caso e allora trovare ciò di cui esso è il numero di Go[“] del.

Il fatto di poter codificare successioni finite di formule ha come conseguenza che possiamo codificare le dimostrazioni nel senso di Hilbert. Abbiamo quindi un modo di descrivere, mediante termini del linguaggio (i numeri) non soltanto il linguaggio stesso, ma anche elementi del metalinguaggio.

Consideriamo ora una funzione $f: \mathbb{N}_n \rightarrow \mathbb{N}$; essa si dirà *rappresentabile* nella teoria

S dei numeri naturali se e soltanto se esiste una formula $A(x_1, \dots x_{n+1})$ con variabili libere $x_1, \dots x_{n+1}$ tale che, comunque assegnati i numeri $k_1, \dots k_{n+1}$, si abbia:

1. se $f(k_1, \dots k_n) = k_{n+1}$ allora $\vdash_S A(\kappa_1, \dots \kappa_{n+1})$, dove $\kappa_1, \dots \kappa_{n+1}$ sono i termini nella teoria che rappresentano rispettivamente i numeri $k_1, \dots k_{n+1}$;
2. $\vdash_S \exists !x_{n+1} A(\kappa_1, \dots, \kappa_n, x_{n+1})$.

Cioè è possibile rappresentare in S mediante una formula il legame funzionale stabilito da f , perchè riusciamo con la seconda condizione ad esprimere anche l'univocità.

Tra le funzioni numeriche hanno evidentemente centrale importanza quelle definite mediante uno *schema ricorsivo*, cioè, sostanzialmente per induzione:

queste funzioni sono infatti effettivamente calcolabili. Abbiamo visto esempi molto semplici nella definizione si somma e prodotto. Si può dimostrare il seguente fatto:

Proposizione 9.2. *Una funzione (parziale) è rappresentabile in S se e soltanto se è ricorsiva.*

Consideriamo ora la funzione $w(x)$ che, calcolata per il numero di Go^{..} del di una formula $B(x)$, dia come risultato il numero di Go^{..} del della dimostrazione di $B(x)$. Si dimostra che $w(x)$ è ricorsiva e, quindi, rappresentabile in S tramite una formula $W(x, x)$ per la quale

1. se $w(u) = v$ allora $\rightarrow_S W(u, v)$
2. se $\neg(w(u) = v)$ allora $\rightarrow_S \neg W(u, v)$

Definiamo allora la formula G così fatta:

$$\forall x' \neg W(x, x').$$

Il significato di questa formula è allora: “La formula codificata dal numero x non ha dimostrazione”. Sia m il numero di Go“ del di G e definiamo una nuova formula G (chiusa) sostituendo m ad x in G :

$$\forall x' \neg W(m, x').$$

Il significato di G è: “La formula G non ha dimostrazione”.

Teorema 9.6 (Go“ del 1931). *Se S è coerente (cioè non ogni formula è dimostrabile) allora in S non si può dimostrare G ; se S è ω -coerente¹ allora in S non si può dimostrare $\neg G$.*

La dimostrazione ripercorre le linee dell’antinomia del mentitore (Epimenide) perchè basata essa pure sull’appiattimento

tra linguaggio e metalinguaggio; l'ipotesi di ω -coerenza è leggermente più forte della coerenza, ma può essere evitata (Rosser 1936) considerando una formula un po' più complessa di G .

Una formula chiusa tale che non esista una dimostrazione nè per lei nè per la sua negazione si dice *indecidibile* e la teoria che ne contenga qualcuna *sintatticamente incompleta*. Aggiungere la formula indecidibile come assioma alla teoria non risolverebbe il problema, anzi, se possibile, lo aggraverebbe, perché avremmo ancora una teoria che contiene S .

La conseguenza più seria (se si può dire così) del Teorema di Gödel è l'indimostrabilità della coerenza di S stessa. Infatti il teorema è formalizzabile in S . Dal momento che la coerenza di S può essere espressa da una formula $\text{Coer}S$

che significhi “Esiste una formula che non è dimostrabile”, allora il Teorema di Gödel si formalizza in S nel seguente modo:

$$\vdash_S \text{Coer}_S \rightarrow G.$$

Se Coer_S fosse dimostrabile, lo sarebbe anche G , usando il Modus Ponens, ma questo non è possibile, perciò le speranze di Hilbert devono essere abbandonate: in una teoria abbastanza potente da contenere l’aritmetica, non saremo mai al riparo da antinomie. Inoltre non tutte le proprietà dei numeri saranno dimostrabili perché avremo sempre formule chiuse indecidibili. In un modello una formula chiusa è vera o falsa, perciò una formula indecidibile indicherà la presenza di una proprietà vera nel modello (quella espressa dalla formula o dalla sua negazione), ma non dimostrabile.

Diamo ora un altro esempio importante

di questione indecidibile che fu proposta nel 1937, insieme alla dimostrazione della sua indecidibilità, dal matematico Alan Turing [54]. Si tratta del *problema della fermata* ossia se sia possibile sempre stabilire la terminazione di un programma per un determinato input finito. È stato dimostrato che non può esistere un algoritmo generale che possa risolvere il problema per tutti i possibili input.

La dimostrazione ricalca da vicino la dimostrazione del Teorema di Gödel. Infatti, un programma può essere assimilato ad una funzione parziale ricorsiva e, quindi, il problema consiste nel decidere, mediante una funzione ricorsiva, se una funzione parziale ricorsiva dà un risultato in corrispondenza ad un certo valore o meno. Come già visto per il caso delle teorie, anche le funzioni (parziali)

ricorsive possono essere codificate attraverso numeri. Supponiamo allora che esista una funzione ricorsiva $g(x, y)$ tale che valga 1 se la funzione fy codificata da y sia definita per x , valga 0 altrimenti. Sia $h(x)$ la funzione che assume valore 1 se $g(x, x) = 0$ e sia indefinita altrimenti; sia io il numero corrispondente ad h . Allora, $h(io) = 1$ se e soltanto se $g(io, io) = 0$. D'altra parte, per come è definita g , se $h(io) = fio$ (io) è definita, allora $g(io, io) = 1$. Abbiamo quindi che $g(io, io) = 1$ se e soltanto se $g(io, io) = 0$, cioè una contraddizione. Ciò vuol dire che non esiste una g siffatta, cioè non esiste una funzione ricorsiva (equiventemente, calcolabile) che stabilisca se una qualunque altra funzione ricorsiva parziale ammetta un valore per un dato input.

Esercizi svolti

Esercizio 9.1. Si formalizzino, usando il linguaggio presentato per il calcolo dei predicati, le seguenti frasi:

- *nessun uomo è mortale e padre di se stesso;*
- *ogni uomo è mortale e non è padre di se stesso;*
- *ogni uomo mortale non è padre di se stesso;*
- *nessun uomo immortale non è padre di se stesso.*

Soluzione: Sia $M(-)$ il predicato “essere mortale” e $P(-, -)$ il predicato “essere padre di”. Allora, le precedenti frasi possono essere formalizzate nel modo seguente:

- $\neg \exists x (M(x) \wedge P(x, x));$

- $\forall x (M(x) \wedge \neg P(x, x));$
- $\forall x (M(x) \rightarrow \neg P(x, x));$
- $\neg \exists x (\neg M(x) \rightarrow \neg P(x, x)).$

Esercizio 9.2. Si formalizzino le seguenti frasi:

- *il Sole è la stella più vicina alla Terra;*
- *i Watussi sono più alti dei Pigmei;*
- *ogni animale ha una madre che è un animale della stessa specie.*

Soluzione:

- Sia $S(-)$ il predicato “essere stella” e $V(-, -)$ il predicato “essere più vicino alla Terra di”; allora, la formula (in questo caso aperta!) cercata è $\forall x (S(x) \rightarrow V(y, x))$, dove y andrà interpretata con la costante *Sol* e (e ciò chiuderà la formula).
- Sia $W(-)$ il predicato “essere un

watusso”, $P(-)$ il predicato “essere un pigmeo” e $A(-, -)$ il predicato “essere più alto di”; allora, la formula cercata è $\forall x \forall y (W(x) \wedge P(y) \rightarrow A(x, y))$ o, equivalentemente, $\forall x (W(x) \rightarrow \forall y (P(y) \rightarrow A(x, y)))$.

- Sia $M(-, -)$ il predicato “essere madre di” e $S(-, -)$ il predicato “essere della stessa specie”; allora, la formula cercata è $\forall x \exists y (M(y, x) \wedge S(x, y))$.

Esercizio 9.3. Si assuma che $P(x)$ denoti la frase “ $x + 2 > 5$ ”.

1. Che tipo di formula è $P(x)$?
2. Dire su quali dei seguenti insiemi può essere definita tale formula:

- (a) \mathbb{N} ;
- (b) $M = \{-1, -2, -3, \dots\}$;
- (c) $O = \{a, b, c, \dots\}$.

- 3. Si discuta la validità/soddisfabilità/contraddittorietà della formula nei precedenti insiemi sui quali è definita.*
- 4. Si dia un insieme sul quale tale formula è una tautologia.*

Soluzione: La formula è una formula predicativa, non quantificata e aperta. Essa è definita soltanto sugli insiemi dei punti (a) e (b); nel primo caso essa risulta soddisfacibile, nel secondo caso contraddittoria. Per rendere la formula una tautologia, basta prendere come insieme di definizione, ad esempio, $\{x \in \mathbb{N} : x > 3\}$.

Esercizio 9.4. Determinare l'insieme di verità di ciascuna delle seguenti formule definite su \mathbb{R} :

1. $\forall x (\mid x \models x)$;

2. $\exists x (x \cdot 2 = x);$
3. $\forall x (x + 1 < x);$
4. $\exists x (x + 2 = x);$
5. $\exists x (|x| = 0).$

Soluzione: Nel primo caso, basta prendere i reali non negativi; nel secondo caso e nell'ultimo caso, si può considerare tutto \mathbb{R} (esiste infatti in \mathbb{R} un numero, in particolare 0, il cui quadrato è uguale a sè e il cui modulo è 0); nel terzo e nel quarto caso, infine, l'insieme di verità è vuoto, poichè nessun numero reale x può soddisfare $x + 1 < x$ o $x + 2 = x$.

Esercizio 9.5. Negare le formule dell'[Esercizio 9.4](#) e trovarne l'insieme di verità.

Soluzione:

1. $\exists x (|x| \neq x);$

2. $\forall x (x^2 \neq x);$
3. $\exists x (x + 1 \geq x);$
4. $\forall x (x + 2 \neq x);$
5. $\forall x (|x| \neq 0).$

Per quanto riguarda i loro insiemi di verità, essi sono, nell'ordine, \mathbb{R} , $\mathbb{R} \setminus \{0, 1\}$, \mathbb{R} , \mathbb{R} , $\mathbb{R} \setminus \{0\}$.

Esercizio 9.6. Sia $A = \{1, 2, 3, 4, 5\}$. Determinare il valore di verità di ciascuna delle seguenti formule definite su A :

1. $\exists x (x + 3 = 10);$
2. $\forall x (x + 3 < 10);$
3. $\exists x (x + 3 < 5);$
4. $\forall x (x + 3 \leq 7).$

Soluzione: Il primo e l'ultimo sono falsi, il secondo e il terzo veri.

Esercizio 9.7. Dato $B = \{2, 3, \dots, 8, 9\}$, trovare un controesempio per ciascuna delle formule seguenti definite su B :

1. $\forall x (x + 5 < 12)$;
2. $\forall x (x \text{ è primo})$;
3. $\forall x (x^2 > 1)$;
4. $\forall x (x \text{ è pari})$.

Soluzione: Nel primo caso, basta prendere $x \in \{7, 8, 9\}$; nel secondo caso, basta prendere $x \in \{4, 6, 8, 9\}$; il terzo caso non ha controesempi (la formula è una tautologia in B); nel quarto caso, basta prendere $x \in \{3, 5, 7, 9\}$.

Esercizio 9.8. Sia $A = \{1, 2, \dots, 9, 10\}$. Si consideri ciascuna delle formule seguenti: se è un enunciato, se ne determini il valore di verità; se è una formula aperta, determinare il suo insieme di verità.

1. $\forall x \exists y(x + y < 14)$;
2. $\forall y(x + y < 14)$;
3. $\forall x \forall y(x + y < 14)$;
4. $\exists y(x + y < 14)$.

Soluzione: La prima è vera; la seconda è una formula aperta con insieme di verità (rispetto ad A) $\{1, 2, 3\}$; la terza è falsa (si prenda, ad esempio, $x = y = 7$); la quarta è aperta e ha come insieme di verità tutto A .

Esercizio 9.9. Sia data la formula $\forall x (P(x) \vee Q(x)) \rightarrow (\forall x P(x) \vee \forall x Q(x))$.

1. La seguente dimostrazione di validità è corretta?

$$\begin{array}{c}
\neg(\forall x(P(x) \vee Q(x)) \rightarrow (\forall x P(x) \vee \forall x Q(x))) \\
| \\
\forall x(P(x) \vee Q(x)), \neg(\forall x P(x) \vee \forall x Q(x)) \\
| \\
\forall x(P(x) \vee Q(x)), \neg\forall x P(x), \neg\forall x Q(x) \\
| \\
\forall x(P(x) \vee Q(x)), \neg\forall x P(x), \neg Q(a) \\
| \\
\forall x(P(x) \vee Q(x)), \neg P(a), \neg Q(a) \\
| \\
P(a) \vee Q(a), \neg P(a), \neg Q(a) \\
/ \quad \backslash \\
P(a), \neg P(a), \neg Q(a) \quad Q(a), \neg P(a), \neg Q(a) \\
\diamond \quad \diamond
\end{array}$$

Se no, correggerla. È possibile dare un tableau chiuso per la negazione della formula in esame?

2. Trovare almeno un modello per la formula.

Soluzione: La costruzione del tableau non è corretta: alla quinta riga si è usata per istanziare $\neg \forall x P(x)$ la costante a già usata alla riga precedente per istanziare $\neg \forall x Q(x)$. Inoltre, alla sesta riga (e seguenti) bisognava mantenere $\forall x (P(x) \vee Q(x))$, pur avendolo istanziato (lecitamente qui) con la costante a . Il

tableau corretto avrebbe, quindi, come nodi $\forall x (P(x) \vee Q(x))$, $P(a)$, $\neg P(b)$, $\neg Q(a)$ e $\forall x (P(x) \vee Q(x))$, $Q(a)$, $\neg P(b)$, $\neg Q(a)$, il secondo dei quali è chiuso. Re sta da capire se il primo nodo genera un tableau chiuso o no; la risposta è negativa. Infatti, per creare una contraddizione con $\neg Q(a)$ dovremmo nuovamente istanziare l'universale con a ; ma questo porterebbe a un nodo $\forall x (P(x) \vee Q(x))$, $P(a)$, $P(a)$, $\neg P(b)$, $\neg Q(a)$ ancora aperto e “dello stesso tipo” del precedente. Per creare una contraddizione con $\neg P(b)$ dovremmo istanziare l'universale con b ; ma questo porterebbe a un nodo $\forall x (P(x) \vee Q(x))$, $Q(b)$, $P(a)$, $\neg P(b)$, $\neg Q(a)$ ancora aperto e riconducibile alla forma del nodo originario.

Per trovare un modello della formula data, ne costruiamo (parzialmente) il tableau (aperto):

$$\begin{array}{c}
 \forall x(P(x) \vee Q(x)) \rightarrow (\forall x P(x) \vee \forall x Q(x)) \\
 \neg(\forall x(P(x) \vee Q(x))) \quad / \quad \backslash \quad \forall x P(x) \vee \forall x Q(x) \\
 \neg(P(a) \vee Q(a)) \\
 \neg P(a), \neg Q(a)
 \end{array}$$

Il ramo di sinistra ci suggerisce come modello $D = \{a\}$ e $|P| = |Q| = \{a\}$.

Esercizio 9.10. *Dimostrare mediante tableau la validità delle formule del tipo seguente:*

$$\begin{aligned}
& \neg(\exists x \forall y_1 \dots \forall y_n P(x, y_1, \dots, y_n) \rightarrow \forall y_1 \dots \forall y_n \exists x P(x, y_1, \dots, y_n)) \\
& \quad | \\
& \quad \exists x \forall y_1 \dots \forall y_n P(x, y_1, \dots, y_n), \neg \forall y_1 \dots \forall y_n \exists x P(x, y_1, \dots, y_n) \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \forall y_1 \dots \forall y_n \exists x P(x, y_1, \dots, y_n) \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \forall y_2 \dots \forall y_n \exists x P(x, b_1, y_2, \dots, y_n) \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \forall y_3 \dots \forall y_n \exists x P(x, b_1, b_2, y_3 \dots, y_n) \\
& \quad | \\
& \quad \dots \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n) \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
& \quad \quad \neg P(a, b_1, b_2, \dots, b_n) \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
& \quad \forall y_2 \dots \forall y_n P(a, b_1, y_2 \dots, y_n), \neg P(a, b_1, b_2, \dots, b_n) \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
& \quad \forall y_3 \dots \forall y_n P(a, b_1, b_2, y_3 \dots, y_n), \neg P(a, b_1, b_2, \dots, b_n) \\
& \quad | \\
& \quad \dots \\
& \quad | \\
& \quad \forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
& \quad P(a, b_1, b_2, \dots, b_n), \neg P(a, b_1, b_2, \dots, b_n) \\
& \quad \diamondsuit
\end{aligned}$$

Soluzione: Costruiamo il tableau della negata:

$$\begin{array}{c}
\neg(\exists x \forall y_1 \dots \forall y_n P(x, y_1, \dots, y_n) \rightarrow \forall y_1 \dots \forall y_n \exists x P(x, y_1, \dots, y_n)) \\
| \\
\exists x \forall y_1 \dots \forall y_n P(x, y_1, \dots, y_n), \neg \forall y_1 \dots \forall y_n \exists x P(x, y_1, \dots, y_n) \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \forall y_1 \dots \forall y_n \exists x P(x, y_1, \dots, y_n) \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \forall y_2 \dots \forall y_n \exists x P(x, b_1, y_2, \dots, y_n) \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \forall y_3 \dots \forall y_n \exists x P(x, b_1, b_2, y_3 \dots, y_n) \\
| \\
\vdots \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n) \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
| \\
\forall y_2 \dots \forall y_n P(a, b_1, y_2 \dots, y_n), \neg P(a, b_1, b_2, \dots, b_n) \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
| \\
\forall y_3 \dots \forall y_n P(a, b_1, b_2, y_3 \dots, y_n), \neg P(a, b_1, b_2, \dots, b_n) \\
| \\
\vdots \\
| \\
\forall y_1 \dots \forall y_n P(a, y_1, \dots, y_n), \neg \exists x P(x, b_1, b_2, \dots, b_n), \\
| \\
P(a, b_1, b_2, \dots, b_n), \neg P(a, b_1, b_2, \dots, b_n) \\
\Diamond
\end{array}$$

Esercizio 9.11. La seguente formula è valida?

$$\begin{array}{c}
\neg(\forall x \exists y P(x, y) \wedge \forall x \neg P(x, y) \wedge \forall x \forall y \forall z (P(x, y) \wedge P(y, z) \rightarrow P(x, z))) \\
/ \qquad | \qquad \backslash \\
\neg \forall x \exists y P(x, y) \qquad \neg \forall x \neg P(x, y) \qquad \neg \forall x \forall y \forall z (P(x, y) \wedge P(y, z) \rightarrow \\
P(x, z))
\end{array}$$

Esiste un modello per essa?

Soluzione: La formula non è valida; infatti, se proviamo a costruirne il tableau della negata otteniamo:

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

/ | \

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

$P(x, z))$

Chiaramente né il ramo di sinistra né quello centrale potranno mai chiudersi, visto che, una volta istanziati opportunamente i quantificatori, conterranno P soltanto in forma o negata (quello di sinistra) o affermata (quello centrale).

Un modello può essere ottenuto costruendo almeno un ramo del tableau della formula data. Un altro modo è quello di andare “a occhio”: se prendiamo $D = \mathbb{N}$, $y = \text{oe } |P| = |<|$, la prima formula della congiunzione ($\forall x \exists y P(x, y)$) dice che ogni numero ha un numero maggiore di

esso (il successore), la seconda formula ($\forall x \neg P(x, y)$) dice che lo o è il minimo rispetto a $<$ e la terza formula dice che $<$ è transitiva.

Esercizio 9.12. *Si dimostri, usando il metodo di Hilbert, che $\rightarrow P(a) \rightarrow \exists x P(x)$.*

Soluzione: Anzitutto, riscriviamo la formula data in termini di implicazioni, negazioni e quantificatori universali; si deve quindi dimostrare $\rightarrow P(a) \rightarrow \neg \forall x \neg P(x)$. A questo punto, la dimostrazione è semplice: basta partire dall'istanza del quarto assioma $\rightarrow \forall x \neg P(x) \rightarrow \neg P(a)$ e applicare le regole di contrapposizione e di doppia negazione viste nel capitolo 8.4.2.

Esercizio 9.13. *Si dimostri, usando il metodo di Hilbert, che $\rightarrow \forall x P(x) \rightarrow \forall y P$*

(y).

Soluzione: La dimostrazione è la seguente:

| | |
|---|-----------|
| $\forall x P(x) \vdash \forall x P(x) \rightarrow P(a)$ | Ax.4 |
| $\forall x P(x) \vdash \forall x P(x)$ | Ip. |
| $\forall x P(x) \vdash P(a)$ | M.P. |
| $\forall x P(x) \vdash \forall y P(y)$ | Gener. |
| $\vdash \forall x P(x) \rightarrow \forall y P(y)$ | Deduzione |

Esercizio 9.14. Si dimostri, usando il metodo di Hilbert, che $\vdash \forall x P(x) \rightarrow \exists x P(x)$.

Soluzione: La dimostrazione è la seguente:

| | |
|---|----------------|
| $\forall x P(x) \vdash \forall x P(x) \rightarrow P(a)$ | Ax.4 |
| $\forall x P(x) \vdash \forall x P(x)$ | Ip. |
| $\forall x P(x) \vdash P(a)$ | M.P. |
| $\forall x P(x) \vdash P(a) \rightarrow \exists x P(x)$ | Esercizio 9.12 |
| $\forall x P(x) \vdash \exists x P(x)$ | M.P. |
| $\vdash \forall x P(x) \rightarrow \exists x P(x)$ | Deduzione |

Esercizio 9.15. Si dimostri, usando il metodo di Hilbert, che $\rightarrow \forall x (P(x) \rightarrow Q)$

$$\rightarrow (\exists x P(x) \rightarrow Q).$$

| | |
|---|-------------------|
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash \forall x(P(x) \rightarrow Q) \rightarrow (P(a) \rightarrow Q)$ | Ax.4 |
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash \forall x(P(x) \rightarrow Q)$ | Ip. |
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash P(a) \rightarrow Q$ | M.P. |
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash \neg Q \rightarrow \neg P(a)$ | Contrapp. |
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash \neg Q$ | Ip. |
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash \neg P(a)$ | M.P. |
| $\forall x(P(x) \rightarrow Q), \neg Q \vdash \forall x \neg P(x)$ | Generalizz. |
| $\forall x(P(x) \rightarrow Q) \vdash \neg Q \rightarrow \forall x \neg P(x)$ | Deduzione |
| $\forall x(P(x) \rightarrow Q) \vdash \neg \forall x \neg P(x) \rightarrow Q$ | Contrapp. |
| $\forall x(P(x) \rightarrow Q) \vdash \exists x P(x) \rightarrow Q$ | Def. di \exists |
| $\vdash \forall x(P(x) \rightarrow Q) \rightarrow (\exists x P(x) \rightarrow Q)$ | Deduzione |

Esercizio 9.16. Si dimostri, usando il metodo di Hilbert, che $\vdash \exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$.

Soluzione: La dimostrazione è la seguente:

| | |
|---|-------------------|
| $\forall y P(a, y) \vdash \forall y P(a, y) \rightarrow P(a, b)$ | Ax.4 |
| $\forall y P(a, y) \vdash \forall y P(a, y)$ | Ip. |
| $\forall y P(a, y) \vdash P(a, b)$ | M.P. |
| $\forall y P(a, y) \vdash P(a, b) \rightarrow \exists x P(x, b)$ | Esercizio 9.12 |
| $\forall y P(a, y) \vdash \exists x P(x, b)$ | M.P. |
| $\vdash \forall y P(a, y) \rightarrow \exists x P(x, b)$ | Deduzione |
| $\vdash \forall y (\forall y P(a, y) \rightarrow \exists x P(x, y))$ | General. |
| $\vdash \forall y (\forall y P(a, y) \rightarrow \exists x P(x, y))$ | |
| $\rightarrow (\forall y P(a, y) \rightarrow \forall y \exists x P(x, y))$ | Ax.5 |
| $\vdash \forall y P(a, y) \rightarrow \forall y \exists x P(x, y)$ | M.P. |
| $\vdash \neg \forall y \exists x P(x, y) \rightarrow \neg \forall y P(a, y)$ | Contrapp. |
| $\vdash \forall x (\neg \forall y \exists x P(x, y) \rightarrow \neg \forall y P(x, y))$ | General. |
| $\vdash \forall x (\neg \forall y \exists x P(x, y) \rightarrow \neg \forall y P(x, y))$ | |
| $\rightarrow (\neg \forall y \exists x P(x, y) \rightarrow \forall x \neg \forall y P(x, y))$ | Ax.5 |
| $\vdash \neg \forall y \exists x P(x, y) \rightarrow \forall x \neg \forall y P(x, y)$ | M.P. |
| $\vdash \neg \forall x \neg \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$ | Contrapp. |
| $\vdash \exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$ | Def. di \exists |

Esercizi da svolgere

Esercizio 9.17. Si formalizzino le seguenti frasi:

- *ogni uomo è più intelligente di ogni animale;*
- *qualche uomo non è più intelligente di ogni animale;*
- *qualche uomo non è più intelligente di*

qualche animale;

- *qualche uomo è più intelligente di ogni animale ma meno forte;*
- *nessun uomo è più intelligente e più forte di ogni animale;*
- *se c'è un uomo più intelligente di ogni animale, allora esiste un animale più forte di quell'uomo.*

Esercizio 9.18. *Si considerino le seguenti formule:*

1. $\forall x P(x) \wedge \exists y Q(y);$
2. $\exists x P(x) \vee \forall y Q(y);$
3. $\exists x \forall y P(x, y);$
4. $\forall x \forall y P(x, y);$
5. $\exists y \exists x \forall z P(x, y, z);$
6. $\forall x \exists y (P(x) \vee Q(y));$
7. $\exists x \forall y (P(x, y) \rightarrow Q(x, y));$
8. $\exists y \exists x (P(x) \wedge \neg Q(y)).$

Per ciascuna di esse, se ne dia un modello (ammesso che sia possibile), la si neghi e si dia un modello (se possibile) per la formula risultante.

Esercizio 9.19. *Dato $\{1, 2, 3\}$ come insieme di definizione, determinare il valore di verità di ciascuna delle formule seguenti e, nel caso in cui sia falso, darne un controesempio:*

1. $\exists x \forall y (x^2 < y + 1);$
2. $\forall x \exists y (x^2 + y^2 < 12);$
3. $\forall x \forall y (x^2 + y^2 < 12);$
4. $\exists x \forall y \exists z (x^2 + y^2 < 2z^2);$
5. $\exists x \exists y \forall z (x^2 + y^2 < 2z^2).$

Esercizio 9.20. *Sia data la formula $\forall x \forall y P(x, y) \rightarrow \exists x \exists y P(x, y)$. Usando il metodo dei tableau, dimostrare che la formula è valida; cosa si può dire*

dell'implicazione inversa? Trovare almeno un modello per essa.

Esercizio 9.21. *La seguente formula è valida?*

$$\begin{aligned} \forall x \exists y P(x, y) \wedge \forall x \neg P(x, y) \wedge \forall x (Q(x) \wedge \neg Q(x)) \wedge \\ \forall x \forall y \forall z (P(x, y) \wedge P(y, z) \rightarrow P(x, z)) \end{aligned}$$

Esiste un modello per essa? Si giustifichi sia intuitivamente che con il metodo dei tableau la risposta precedente.

Esercizio 9.22. *Sia data la formula aperta $A(x, y)$ e l'interpretazione $|A| = |“Ama”|$ e D è un insieme non vuoto di persone.*

1. *Che differenze ci sono fra le seguenti formule?*

- (a) $\forall x \exists y A(x, y);$
- (b) $\exists y \forall x A(x, y);$
- (c) $\exists x \forall y A(x, y);$

(d) $\forall y \exists x A(x, y)$.

(Sugg.: esprimere le formule in linguaggio comune).

2. Usando il metodo dei tableau, si dica quali fra le seguenti formule sono valide:

(a) $\exists y \forall x A(x, y) \rightarrow \forall x \exists y A(x, y)$;

(b) $\forall x \exists y A(x, y) \rightarrow \exists x \forall y A(x, y)$;

(c) $\exists x \forall y A(x, y) \rightarrow \forall y \exists x A(x, y)$;

(d) $\forall y \exists x A(x, y) \rightarrow \exists x \forall y A(x, y)$.

Esercizio 9.23. Si svolgano gli Esercizi 9.10, 9.11, 9.20, 9.21 e 9.22 con il metodo di Gentzen.

Esercizio 9.24. Si dimostri, usando il metodo di Hilbert, che $\vdash (P \rightarrow \forall x Q(x)) \rightarrow \forall x (P \rightarrow Q(x))$.

Esercizio 9.25. Si dimostri, usando il metodo di Hilbert, che $\vdash (\exists x P(x) \rightarrow Q) \rightarrow \forall x (P(x) \rightarrow Q)$.

Esercizio 9.26. Si dimostri, usando il metodo di Hilbert, che $\vdash \exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$.

¹ Una teoria T è *ω -coerente* se potendo dimostrare in essa una proprietà $P(x)$ per qualunque numero naturale, non si può dimostrare anche $\exists x \neg P(x)$

Procedimenti di risoluzione

Essendo questo volume dedicato prevalentemente a futuri informatici, riveste particolare interesse il problema dell'automatizzabilità delle procedure di deduzione. È abbastanza evidente che non si può nemmeno sperare di automatizzare una procedura di deduzione nello stile di Hilbert, perché quel tipo di dimostrazione prevede un ragionamento euristico molto sofisticato. Maggiori possibilità sembra promettere un metodo meccanico come quello dei tableau. Nel caso del calcolo dei predicati, però, non è detto che tale

metodo dia sempre una risposta in tempo finito, perchè, se il tableau si chiude, esso si chiude dopo un numero finito di passi, ma, se è destinato a rimanere aperto, potremmo andare avanti senza esserne mai sicuri. Ciò è dovuto alla presenza delle γ -regole.

Da quella esperienza abbiamo tuttavia imparato che un buon candidato a costituire una procedura automatizzabile dovrebbe essere un procedimento per *refutazione*, cioè un procedimento che, come nel caso dei tableau, vada alla ricerca di una contraddizione producendo il controesempio. Dovremo poi preoccuparci di rendere efficiente il metodo. Faremo quindi in quest'ultimo capitolo i primi passi per acquisire il metodo detto di *risoluzione*. Come di consueto, impareremo il metodo innanzitutto per il calcolo degli enunciati e

lo estenderemo poi con i quantificatori. L’idea di base consiste nel ridurre una formula ad un insieme di letterali la cui mutua (in)soddisfacibilità sia immediatamente verificabile, o quasi. Nell’estensione al calcolo dei predicati, dovremo trovare un modo di eliminare anche i quantificatori (cosa che non era possibile con i tableau).

Questo secondo passo costituirà un’occasione per occuparci di una classe di termini del linguaggio, che, per semplicità, avevamo volutamente trascurato nei paragrafi precedenti: i termini definiti attraverso funzioni. Infatti, nella ricerca di un controesempio l’uso di variabili e costanti soltanto porterebbe ad una rigidità inopportuna. I simboli funzionali, infatti, ci permettono di esprimere la dipendenza di una variabile da un’altra senza usare i

quantificatori. Si immagini ad esempio la formula

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

Ora x rappresenta un qualunque individuo nel dominio di interpretazione, mentre y non è qualunque e nemmeno è una specifica costante. Lo possiamo pensare come un individuo che va scelto in base ad x , cioè che *dipende* da x .

Possiamo esprimere questa dipendenza mediante un simbolo funzionale f , dicendo che y è $f(x)$. Certamente in questo modo forziamo un po' la situazione, perchè non è proprio detto che il legame tra x ed y sia di tipo funzionale. Stiamo di fatto imponendo una condizione più rigida di quella definita dai quantificatori, perchè diciamo in questo modo che y è *univocamente* determinato da x (mentre, in generale, potrebbero esistere diversi

elementi del dominio in grado di soddisfare $P(x, -)$). D'altra parte, essendo noi alla ricerca di un controesempio, se lo trovassimo anche in questa situazione più ristretta, saremmo sicuri che esso varrebbe nella situazione generale.

10.1 La risoluzione per il calcolo degli enunciati

10.1.1 Forme clausali

Cominciamo con il rilevare che ogni formula del calcolo degli enunciati può essere trasformata in una formula (logicamente) equivalente che assume una forma particolare, detta *forma normale congiuntiva*. Tale formato risulterà molto utile per l'impostazione del procedimento risolutivo.

Definizione 10.1. Una formula è in forma normale congiuntiva se è costituita da una congiunzione di disgiunzioni di enunciati atomici (eventualmente negati), chiamati letterali. Ognuno dei blocchi di disgiunzione viene detto clausola.

Esempio 10.1. La formula $(P \vee Q \vee (\neg Q)) \wedge ((\neg P) \vee (\neg Q) \vee R)$ è in forma normale congiuntiva.

Esempio 10.2. (Controesempio) La formula $P \vee (Q \wedge (P \rightarrow R))$ non è in forma normale congiuntiva. Essa può però essere trasformata in una formula in forma normale congiuntiva ricordando che:

$A \vee (B \wedge C)$ equivale a $(A \vee B) \wedge (A \vee C)$

$A \rightarrow B$ equivale a $(\neg A) \vee B$

Applicando la prima osservazione, la

formula data può essere riscritta in $(P \vee Q) \wedge (P \vee (P \rightarrow R))$ e, applicando la seconda osservazione, in $(P \vee Q) \wedge (P \vee (\neg P) \vee R)$. Quest'ultima formula è in forma normale congiuntiva.

Proposizione 10.1. *Un enunciato in forma normale congiuntiva è soddisfacibile se e soltanto se possiamo contemporaneamente soddisfare almeno un letterale in ogni termine della congiunzione (clausola).*

Dimostrazione. La dimostrazione segue immediatamente dalle tavole di verità dei connettivi \wedge e \vee e viene lasciata come esercizio.

Si può anche facilmente dimostrare che un enunciato in forma normale congiuntiva è una tautologia se e soltanto se in ogni clausola è presente (almeno) un

enunciato con la sua negazione.

Nota 10.1. *A volte possono essere considerati anche enunciati in forma normale disgiuntiva, costituiti da una disgiunzione di congiunzioni di enunciati atomici (eventualmente negati). Si può dimostrare che un enunciato in forma normale disgiuntiva è insoddisfacibile se e solo se in ogni blocco di disgiunzioni è presente (almeno) un enunciato con la sua negazione. Nel presente capitolo ci occuperemo però solamente di enunciati in forma normale congiuntiva.*

Sarà essenziale, per rendere possibile l'applicazione del metodo che illustreremo, che le formule di volta in volta considerate siano poste in forma normale congiuntiva. Ciò è possibile applicando la procedura seguente:

1. si scriva innanzitutto la formula data eliminando i connettivi presenti eccetto \wedge e \vee (si utilizzino le equivalenze indicate nell'[Esempio 8.7](#));
2. si spostino tutte le negazioni che operano su parentesi all'interno di tali parentesi utilizzando le leggi di De Morgan;
3. eliminare eventuali doppie negazioni;
4. eliminare infine eventuali congiunzioni presenti all'interno delle parentesi utilizzando le leggi distributive:
 $A \vee (B \wedge C)$ equivale a $(A \vee B) \wedge (A \vee C)$
 $(A \wedge B) \vee C$ equivale a $(A \vee C) \wedge (B \vee C)$
5. eliminare eventuali occorrenze multiple dello stesso letterale, ricordando che:
 $A \wedge A$ equivale a A
 $A \vee A$ equivale a A

Esempio 10.3. La seguente successione di formule mostra i cinque passi applicati alla formula $(\neg p \Rightarrow \neg q) \Rightarrow (p \Rightarrow q)$:

1. $\neg(\neg\neg p \vee \neg q) \vee (\neg p \vee q)$
2. $(\neg\neg\neg p \wedge \neg\neg q) \vee (\neg p \vee q)$
3. $(\neg p \wedge q) \vee (\neg p \vee q)$
4. $(\neg p \vee q \vee \neg p) \wedge (\neg p \vee q \vee q)$
5. $(\neg p \vee q \vee \neg p) \wedge (\neg p \vee q)$

Una clausola sarà rappresentata da un insieme (finito) di letterali da intendersi disgiunti tra loro; una clausola costituita da un solo letterale si dice *unitaria*. Una formula (in forma normale congiuntiva) *in forma clausale* è scritta come un insieme di clausole da intendersi congiunte tra loro.

I letterali di una clausola si scrivono successivamente, senza alcun simbolo, e le negazioni si indicano sottolineando il

letterale. Se l è un letterale, indichiamo con l^c il suo complemento (quindi se l è P , l^c è \underline{P} e viceversa). Le clausole si separano tra loro mediante una virgola che sostituisce la congiunzione. In questo modo l'enunciato iniziale si presenta come un insieme (finito) di insiemi (finiti) di letterali. Per quanto detto nella Proposizione 10.1, l'enunciato originale sarà soddisfacibile se e soltanto se riusciremo a soddisfare contemporaneamente un letterale per ogni clausola. Di conseguenza l'insieme vuoto di clausole (in simboli $\{\}$) sarà sempre soddisfacibile, mentre la clausola vuota (in simboli $[]$) non lo sarà mai e rappresenterà, quindi, la contraddizione.

Esempio 10.4. La formula in forma normale congiuntiva $(P \vee Q \vee (\neg Q)) \wedge ((\neg P) \vee (\neg Q) \vee R)$ può esprimersi in forma clausale S nel modo seguente:

$$S = \{P\cancel{Q}Q, \cancel{P}QR\}$$

10.1.2 Risoluzione

In una procedura per refutazione ci viene chiesto di stabilire la soddisfacibilità o meno di un insieme di formule.

Considerate due formule in forma clausale S e T , scriveremo $S \approx T$ quando S è soddisfacibile se e soltanto se T è soddisfacibile.

Passo di risoluzione. Siano C e D due clausole tali che $l \in C$ e $l^c \in D$ (esse vengono dette *clausole contrastanti* sui letterali complementari l e l^c); allora diremo *risolvente* di C e D la clausola $\text{Res}[C, D]$ contenente tutti i letterali di C tranne l e tutti i letterali di D tranne l^c . C e D sono anche dette *genitrici* di $\text{Res}[C, D]$.

Teorema 10.1. *La risolvente $\text{Res}[C, D]$ è soddisfacibile se e solo se le genitrici C e D*

sono soddisfacibili.

Dimostrazione. Se $\text{Res}[C, D]$ è soddisfacibile, possiamo trovare un'assegnazione di valori di verità agli atomi che vi compaiono in modo che almeno un suo letterale sia vero; se questo letterale apparteneva originariamente, diciamo a C , allora anche C è soddisfacibile con la stessa assegnazione. D'altra parte questo letterale non poteva essere né l né l^c , quindi possiamo estendere l'assegnazione in modo che l^c risulti vero e così sia soddisfatta anche D . Viceversa, se C e D sono entrambe soddisfatte, vuol dire che esistono due letterali, uno in C ed uno in D veri per una certa assegnazione. Questi due letterali non possono essere l ed l^c perché contrastanti, quindi esisterà un letterale diverso da essi che risulterà vero. Questo letterale sarà presente anche in $\text{Res}[C, D]$.

Esempio 10.5. Si considerino le clausole $P Q$ e $\underline{P} \underline{Q}$. Esse ammettono due risolventi: se si sceglie di risolvere su P , si ottiene la risolvenete $Q \underline{Q}$; se si sceglie di risolvere su Q , si ottiene la risolvenete $P \underline{P}$.

Entrambe le risolventi sono banalmente soddisfacibili (anzi, sono tautologie). Per il Teorema 10.1, anche le due clausole iniziali sono soddisfacibili: basta per esempio assegnare “vero” a P e “falso” a Q . Si noti invece che, sebbene le risolventi siano tautologie, la congiunzione delle due clausole iniziali non lo è: basta assegnare a P e Q lo stesso valore di verità. Quindi, il Teorema 10.1 può soltanto essere espresso in termini di soddisfacibilità e non di validità.

Procedura risolutiva. Sia S una data formula in forma clausale, cioè un insieme di clausole. Dovremo costruire una sequenza $S_0 \dots S_n$ di formule in forma

clausale, partendo da $S_0 = S$, costruendo ripetutamente la risolvente di due clausole contrastanti.

Consideriamo la formula S_i ; scegliamo due clausole contrastanti C e D di S_i e consideriamo la risolvente $\text{Res}[C, D]$. Costruiamo allora:

$$S_{i+1} = S_i \cup \text{Res}[C, D]$$

Se $\text{Res}[C, D]$ è la clausola vuota, indicata con $[]$, allora la procedura termina e S è insoddisfacibile. Se $S_{i+1} = S_i$ per ogni scelta di clausole contrastanti, allora la procedura termina e S è soddisfacibile. Altrimenti si itera con la costruzione di S_{i+2} .

Osserviamo che la procedura risolutiva ora presentata è, propriamente, una procedura di refutazione: essa può infatti

essere applicata per dimostrare che una formula A è una tautologia grazie al fatto che l'insoddisfacibilità di $\neg A$ corrisponde alla validità di A . Applichiamo il metodo all'esempio seguente.

Esempio 10.6. *Mostriamo la validità di (un'istanza del) secondo assioma di Hilbert. A tale scopo, costruiamo una risoluzione per*

$$\neg((P \rightarrow (Q \rightarrow R)) \rightarrow ((P \rightarrow Q) \rightarrow (P \rightarrow R)))$$

Dapprima dobbiamo portare tale formula in forma clausale; seguendo il procedimento descritto in Sezione 10.1.1 otteniamo

$$S_0 = \{P, \underline{P}Q, \underline{R}, \underline{P}\underline{Q}R\}$$

Prendiamo $C_1 = \{\underline{R}\}$ e $D_1 = \{ \underline{P} \underline{Q} R\}$; tali clausole sono contrastanti sui letterali

complementari R e \underline{R} . La loro risolvente è $\text{Res}[C_3, C_4] = \{\underline{P} \underline{Q}\}$; pertanto,

$$S_1 = \{P, \underline{P} Q, \underline{R}, \underline{P} \underline{Q} R, \underline{P} \underline{Q}\}$$

Prendiamo ora $C_2 = \underline{P} \underline{Q}$ e $D_2 = \underline{P} \underline{Q}$; esse sono contrastanti sui letterali complementari Q e \underline{Q} . La loro risolvente è $\text{Res}[C_2, C_5] = \{\underline{P}\}$; pertanro,

$$S_2 = \{P, \underline{P} Q, \underline{R}, \underline{P} \underline{Q} R, \underline{P} \underline{Q}, \underline{P}\}$$

Infine prendiamo $C_3 = \underline{P}$ e $D_3 = P$; esse sono contrastanti sui letterali complementari P e \underline{P} . La loro risolvente è $\text{Res}[C_1, C_6] = []$ e concludere che l'assegnata $S = \{P, \underline{P} Q, \underline{R}, \underline{P} \underline{Q} R\}$ è insoddisfacibile. Quindi, la formula originaria (cioè, $\neg((P \rightarrow (Q \rightarrow R)) \rightarrow ((P \rightarrow Q) \rightarrow (P \rightarrow R)))$) è insoddisfacibile, da

cui la validità del secondo assioma di Hilbert.

Anche la procedura ora presentata è corretta e completa rispetto alle tavole di verità. Si possono infatti dimostrare i teoremi seguenti:

Teorema 10.2. (*Correttezza della risoluzione*) *Se dalla formula in forma clausale S viene derivata mediante la procedura di risoluzione la clausola insoddisfacibile $[]$, allora S è insoddisfacibile.* *Dimostrazione.* La dimostrazione si fa nel modo consueto per induzione sul numero dei passi di risoluzione, usando il teorema 10.1.

Teorema 10.3. (*Completezza della risoluzione*) *Se la formula in forma clausale S è insoddisfacibile, allora la clausola insoddisfacibile $[]$ viene derivata*

mediante la procedura di risoluzione.

La dimostrazione del teorema di completezza non è affatto semplice e richiede strumenti che non abbiamo introdotto in questo testo.

10.2 La risoluzione per il calcolo dei predicati

La procedura presentata nel caso di formule proposizionali può essere estesa a formule predicative. Tale estensione risulta praticamente utile in quanto altri metodi applicabili al calcolo dei predicati, come ad esempio il metodo dei tableau semantici, risultano di scarsa efficienza.

10.2.1 Forma prenessa e forma clausale

Anche nel caso di formule appartenenti al linguaggio del calcolo dei predicati, vogliamo ottenere una loro trasformazione in insieme finito di clausole in modo da poter estendere facilmente ad esse il metodo di risoluzione. Opereremo nel seguente modo: con un'opportuna ridenominazione delle variabili vincolate faremo in modo che una formula F sia trasformata in un'altra logicamente equivalente F' , ma fatta in modo che i quantificatori compaiano tutti all'inizio; tale formato viene chiamato *forma prenessa*. Poi cancelleremo tutti i quantificatori stando attenti, però a che le variabili quantificate esistenzialmente siano sostituite con termini funzionali che indichino la dipendenza dalle variabili quantificate universalmente che le precedono. Come accennavamo nell'introduzione a questo capitolo, anche se le variabili non più

quantificate saranno intese come quantificate universalmente, la formula così ottenuta F'' non sarà più logicamente equivalente a quella iniziale, ma soltanto *equisoddisfacibile*, cioè $F \approx F''$. Essendo noi interessati alla soddisficiabilità, questo fatto non costituirà un problema e potremo tranquillamente trasformare F in forma clausale, anche se questa volta i letterali saranno costituiti da formule atomiche e negazioni di formule atomiche.

Ricordando che, operando in termini di refutazione, siamo interessati soltanto a trovare istanze delle nostre formule che producano una contraddizione (per arrivare a dimostrarne l'insoddisficiabilità), nel caso di formule predicative possiamo spesso fare un ulteriore passo rispetto a ciò che possiamo fare negli enunciati. Infatti arrivati alla

forma clausale, può accadere che tutti i letterali risultino formule chiuse (senza variabili) e, quindi, possiamo operare esattamente come nel caso degli enunciati, oppure compaiano in essi delle variabili. Basterà allora trovare delle istanze di quegli atomi che permettano di arrivare ad una contraddizione, anche se gli atomi originali non la mostravano direttamente. Si tratta, cioè, se possibile, di sostituire le variabili con opportuni termini in modo da rendere contrastanti clausole che originariamente non lo erano. Questa procedura si chiama *unificazione* ed è l'aspetto più delicato ed interessante della procedura di risoluzione. Si noti che, come nel caso della definizione di un'interpretazione, il passaggio dal calcolo degli enunciati a quello dei predicati richiede un esame dei termini all'interno delle formule, che non possono più essere pensate come scatole chiuse.

Definizione 10.2. Una formula si dice

- *in forma normale congiuntiva prenessa se e soltanto se è espressa come*

$$Q_1 x_1 \dots Q_n x_n P$$

dove $Q_1 \dots Q_n$ sono quantificatori e P una formula priva di quantificatori in forma normale congiuntiva.

- *in forma clausale se è una formula in forma normale congiuntiva prenessa in cui tutti i quantificatori $Q_1 \dots Q_n$ sono universali.*

Proposizione 10.2. Una formula F può essere sempre trasformata in una F' equivalente ad essa ed in forma normale congiuntiva prenessa

Dimostrazione. (cenno) Innanzitutto, si

danno nomi nuovi alle variabili quantificate; poi si sostituiscono tutti i connettivi diversi da \wedge e \vee usando espressioni logicamente equivalenti. Si portano all'interno tutte le negazioni, usando ancora equivalenze logiche. Dopodichè i quantificatori possono essere estratti e posti in testa alla formula rispettando le precedenze soltanto nel caso in cui uno di essi capiti nell'ambito di un altro. Sappiamo già di poter cambiare il nome delle variabili vincolate mantenendo l'equivalenza logica. Il fatto di aver posto i quantificatori all'inizio non altera nulla, perchè, avendo usato nomi nuovi, il loro campo di azione non cambia. La parte che non contiene quantificatori può poi essere messa in forma normale congiuntiva con il metodo illustrato sopra.

La procedura per portare una qualsiasi formula in forma clausale è la seguente:

- Grazie all'[Esercizio 9.13](#), cambiamo il nome alle variabili in modo che nessuna variabile appaia in due quantificatori distinti.
- Eliminiamo tutti i connettivi a parte ‘ \vee ’ e ‘ \wedge ’.
- Spingiamo la negazione all’interno, eliminando eventuali doppie negazioni.

Quando ciò avviene attraverso un quantificatore, usiamo le equivalenze:

$$\begin{aligned}\neg \forall x A(x) &\leftrightarrow \exists x \neg A(x) \\ \neg \exists x A(x) &\leftrightarrow \forall x \neg A(x)\end{aligned}$$

- Estraiamo tutti i quantificatori: ripetutamente scegliamo un quantificatore che non è nell’ambito di un altro quantificatore non ancora estratto, e lo estraiamo usando le regole seguenti (che si possono applicare poichè nessuna variabile appare in due

quantificatori):

$$A \ op \ Q \ x \ B(x) \leftrightarrow Qx \ (A \ op \ B(x))$$

$$Qx \ A(x) \ op \ B \leftrightarrow Qx \ (A(x) \ op \ B)$$

dove Q è un quantificatore e op è ‘ \vee ’ oppure ‘ \wedge ’.

- Usiamo le leggi distributive per trasformare la formula quantificata in forma normale congiuntiva.
- Ora che abbiamo trasformato la formula in forma normale congiuntiva prenessa, il solo passo che ci resta consiste nell'eliminare i quantificatori esistenziali dal prefisso. Sia $\exists x$ un quantificatore esistenziale in A , siano y_1, \dots, y_n le variabili universali quantificate *che precedono* $\exists x$ e sia f un *nuovo* simbolo di funzione n -ario. Cancelliamo $\exists x$ e sostituiamo ogni occorrenza di x con $f(y_1, \dots, y_n)$. Se non ci sono

quantificatori universali che precedono $\exists x$, sostituiamo x con un nuovo simbolo di costante (ossia una funzione di arità nulla) a . Questi nuovi simboli di funzione sono chiamati *funzioni di Skolem*.

Esempio 10.7. Portiamo la formula

$\forall x (p(x) \Rightarrow q(x)) \Rightarrow (\forall x p(x) \Rightarrow \forall x q(x))$
in forma clausale.

- Iniziamo col ridenominare le variabili legate:

$\forall x (p(x) \Rightarrow q(x)) \Rightarrow (\forall y p(y) \Rightarrow \forall z q(z))$

- Eliminiamo tutti i connettivi a parte ‘ \vee ’ e ‘ \wedge ’:

$\neg \forall x (\neg p(x) \vee q(x)) \vee \neg \forall y p(y) \vee \forall z q(z)$

- *Spingiamo la negazione all'interno, eliminando eventuali doppie negazioni:*

$$\exists x (p(x) \wedge \neg q(x)) \vee \exists y \neg p(y) \vee \forall z q(z)$$

- *Estraiamo tutti i quantificatori:*

$$\exists x \exists y \forall z ((p(x) \wedge \neg q(x)) \vee \neg p(y) \vee q(z))$$

- *Usiamo le leggi distributive per trasformare la formula quantificata in forma normale congiuntiva, e otteniamo:*

$$\exists x \exists y \forall z ((p(x) \vee \neg p(y) \vee q(z)) \wedge (\neg q(x) \vee \neg p(y) \vee q(z)))$$

- *Skolemizziamo la forma normale congiuntiva prenessa:*

$$\forall z ((p(a) \vee \neg p(b) \vee q(z)) \wedge (\neg q(a) \vee \neg$$

$$p(b) \vee q(z)))$$

dove a e b sono delle funzioni di Skolem (costanti) corrispondenti rispettivamente alle variabili quantificate esistenzialmente x e y .

Nota 10.2. *Osserviamo che l'ordine di estrazione dei quantificatori è arbitrario. Nell'Esempio 10.7 si sarebbe anche potuta produrre la formula:*

$$\forall z \exists x \exists y ((p(x) \vee \neg p(y) \vee q(z)) \wedge (\neg q(x) \vee \neg p(y) \vee q(z)))$$

Se avessimo scelto questo ordine di estrazione del quantificatore, la forma clausale sarebbe stata:

$$\forall z ((p(f(z)) \vee \neg p(g(z)) \vee q(z)) \wedge (\neg q(f(z)) \vee \neg p(g(z)) \vee q(z)))$$

dove le funzioni di Skolem f e g sono

unarie perchè l'unico quantificatore universale su z precede i quantificatori esistenziali.

Esempio 10.8. Portiamo la formula

$$\exists x \forall y p(x, y) \Rightarrow \forall y \exists x p(x, y)$$

in forma clausale.

- Cambiamo il nome alle variabili vincolate

$$\exists x \forall y p(x, y) \Rightarrow \forall w \exists z p(z, w)$$

- Eliminiamo i connettivi booleani e spingiamo verso l'interno la negazione

$$\forall x \exists y \neg p(x, y) \vee \forall w \exists z p(z, w)$$

- Estraiamo i quantificatori

$$\forall x \exists y \forall w \exists z (\neg p(x, y) \vee p(z, w))$$

- *Sostituiamo i quantificatori esistenziali con le funzioni di Skolem*

$$\forall x \forall w (\neg p(x, f(x)) \vee p(g(x, w), w))$$

Nota 10.3. *In pratica è più conveniente usare una trasformazione leggermente diversa di una formula in forma clausale. Prima di estrarre i quantificatori, spingiamoli all'interno il più possibile; quindi sostituiamo i quantificatori esistenziali con le funzioni di Skolem; infine estraiamo tutti i quantificatori (universali) rimanenti. Questo assicura che il numero dei quantificatori universali che precedono un esistenziale sia minimo, e lo stesso vale per l'arità delle funzioni di Skolem. Nell'Esempio 10.8 otterremmo:*

$$\forall x \neg p(x, f(x)) \vee \forall w p(g(w), w)$$

rimpiazzando la funzione binaria g con

una funzione unaria, da cui

$$\forall x \forall w (\neg p(x, f(x)) \vee p(g(w), w))$$

Vogliamo adesso trasformare la F' in forma clausale F'' in modo che $F' \approx F''$. Quest'ultimo passo ha bisogno del seguente essenziale risultato (dovuto a T. Skolem):

Teorema 10.4. *Sia F una formula chiusa. Allora esiste una formula G in forma clausale tale che F è soddisfacibile se e soltanto se G è soddisfacibile.*

Dimostrazione. (cenno) Come già sappiamo, F può essere messa in forma prenessa normale congiuntiva F' equivalente ad F . Cancelliamo allora dalla F' tutti i quantificatori esistenziali, introducendo per ognuno di essi nel linguaggio un nuovo simbolo funzionale f

$(-, -, \dots, -)$ ad n posti, se n era il numero dei quantificatori universali che precedevano in F' il quantificatore soppresso. Allora nella nuova formula la variabile x originariamente vincolata dal quantificatore soppresso viene sostituita con il termine $f(x_1, \dots, x_n)$, dove x_1, \dots, x_n sono le variabili vincolate dai quantificatori universali che lo precedevano. Questa procedura non rispetta l'equivalenza logica, ma soltanto l'equisoddisfacibilità. Supponiamo infatti di avere un modello M per F' con dominio di interpretazione D . Ciò vuol dire che per ogni assegnazione di valori alle variabili x_1, \dots, x_n in D , possiamo trovare un'assegnazione per la variabile x in modo da soddisfare F' . Scegliendo di volta in volta un'assegnazione possibile, riusciamo a definire una funzione $\varphi: D_n \rightarrow D$. Allora φ può costituire

l'interpretazione di un simbolo funzionale n-ario f in modo tale che sia soddisfatta la nuova formula F'' ottenuta da F' sopprimendo il quantificatore esistenziale. Viceversa se tale formula è soddisfacibile, cioè esiste un modello per essa, allora è vero che comunque assegnati valori alle variabili x_1, \dots, x_n , esiste un'assegnazione per x in modo tale che F' sia soddisfacibile. Basterà infatti assegnare ad x l'interpretazione di $f(x_1, \dots, x_n)$.

Ripetendo questo procedimento per tutti i quantificatori esistenziali, si ottiene una G tale che $F \approx G$.

Una volta arrivati ad avere soltanto quantificatori universali, possiamo toglierli, convenendo che ogni variabile è da intendersi come quantificata universalmente e procedere ottenendo la forma clausale, che sarà formata da un insieme (finito) di insiemi (finiti) di

letterali: in questo caso formule atomiche o negazione di formule atomiche. Se in questa forma non compare nessuna variabile, possiamo procedere con la risoluzione come nel caso del calcolo degli enunciati, altrimenti possiamo tentare di portare due clausole a contrastare anche se non lo possono immediatamente, istanziando opportunamente le variabili.

10.2.2 Unificazione

Siano date dunque due clausole non immediatamente contrastanti contenenti due letterali opposti con lo stesso predicato e con variabili. Si ricordi che stiamo cercando di arrivare ad una contraddizione con un controesempio e che le variabili sono da intendersi quantificate universalmente, perciò istanziabili a piacimento. Attenzione, però, che la sostituzione di una variabile

con un termine andrà fatta in modo uniforme su entrambe le clausole che vogliamo portare a contrastare.

Esempio 10.9. Consideriamo i due letterali $P(f(x), g(y))$ e $\neg P(ff(a), g(z))$. Questi non sono immediatamente contrastanti, ma lo sarebbero se riuscissimo a risolvere qualcosa di molto simile ad un sistema di equazioni, cioè

$$\begin{aligned}f(x) &= ff(a) \\g(y) &= g(z)\end{aligned}$$

In questo caso molto semplice si vede immediatamente quale sostituzione dobbiamo fare per “risolvere” il nostro sistema. Ad esempio, potremmo sostituire x, y e z con $f(a)$; ma basterebbe anche sostituire x con $f(a)$ e y e z con w o, ancora più semplicemente, x con $f(a)$ e y con z .

Una sostituzione può essere rappresentata da qualcosa che ha la forma di un sistema di equazioni, questa volta in forma esplicita, anche se l'uguaglianza deve ritenersi orientata, perché il termine di destra va sostituito a quello di sinistra e non viceversa. Le tre sostituzioni che abbiamo descritto avrebbero potuto anche scriversi

1. $x = f(a), y = f(a), z = f(a);$
2. $x = f(a), y = w, z = w;$
3. $x = f(a), y = z.$

Alcune sostituzioni sono più generali di altre, nel senso che quella più particolare si può ottenere da quella più generale mediante un'ulteriore istanziazione.

Nell'esempio precedente la prima può ottenersi dalla seconda facendo seguire quest'ultima da $w = f(a)$. È chiaro che, nel tentativo di porre a contrasto due clausole,

ci interesseranno le sostituzioni più generali, perchè dovendo probabilmente riapplicare il metodo, non vogliamo aver fatto sostituzioni inutilmente troppo vincolanti (ad esempio introducendo costanti o simboli funzionali non indispensabili, che poi non potremmo più sostituire). Ci poniamo quindi il problema di trovare, se esiste, l'*unificatore più generale* (di solito abbreviato in *mgu*, dall'inglese “most general unifier”), cioè la sostituzione più generale che porta due clausole a contrastare.

Questo risultato viene raggiunto con un algoritmo (dovuto a Martelli e Montanari [38] e diverso da quello originale dovuto a J.A. Robinson [49]) che opera nel seguente modo. Dato un sistema di equazioni fra termini, operare finchè possibile i seguenti passi:

1. Se si ha l'equazione $t = x$ e t non è una

variabile, trasformarla in $x = t$

2. Cancellare ogni equazione del tipo $x = x$
3. Se $t' = t''$ e sia t' che t'' non sono variabili, ma t' è del tipo $f(t'_1, \dots, t'_k)$ e t' è del tipo $g(t''_1, \dots, t''_k)$, allora, se $f \neq g$, l'algoritmo termina e non esiste unificatore; oppure $f = g$ e l'equazione originaria viene rimpiazzata con le k equazioni

$$t'_1 = t''_1 \quad \dots \quad t'_k = t''_k$$

4. Se $x = t$ è tale che x compaia altrove nel sistema di equazioni, se x compare in t , allora l'algoritmo termina e non esiste unificatore, altrimenti tutte le occorrenze di x vengono sostituite con t .

Prima di dimostrare che questo algoritmo costruisce una mgu, lavoriamo con un esempio.

Esempio 10.10. Proviamo a unificare $P(f(x, g(y), h(z)), h(h(z)))$ e $P(f(h(w), g(x), h(w)), h(y))$. Applicando la regola 3, otteniamo due equazioni:

$$f(x, g(y), h(z)) = f(h(w), g(x), h(w))$$
$$h(h(z)) = h(y)$$

Applichiamo nuovamente la regola 3 a entrambe le equazioni:

$$x = h(w)$$

$$g(y) = g(x)$$

$$h(z) = h(w)$$

$$h(z) = y$$

Applichiamo la regola 1 all'ultima equazione:

$$x = h(w)$$

$$g(y) = g(x)$$

$$h(z) = h(w)$$

$$y = h(z)$$

Applichiamo la regola 3 alla seconda e terza equazione:

$$x = h(w)$$

$$y = x$$

$$z = w$$

$$y = h(z)$$

Usiamo la regola 4 sulla prima e seconda equazione per sostituire le occorrenze di x con $h(w)$, e sulla terza e quarta per sostituire le occorrenze di z con w :

$$x = h(w)$$

$$y = h(w)$$

$$z = a$$

$$y = h(w)$$

Usiamo la regola 4 sulla seconda e quarta equazione per sostituire le occorrenze di y con $h(w)$:

$$x = h(w)$$

$$y = h(w)$$

$$z = a$$

$$h(w) = h(w)$$

Usiamo la regola 3 sull'ultima equazione e poi la regola 2 per cancellare $w = w$; otteniamo

$$x = h(w)$$

$$y = h(w)$$

$$z = w$$

Affermiamo che

$$\{x \leftarrow h(w), y \leftarrow h(w), z \leftarrow w\}$$

è un unificatore più generale. Si noti che la sostituzione, di fatto, unifica i due termini originali; inoltre, dato un altro unificatore come

$$\{x \leftarrow h(f(a)), y \leftarrow h(f(a)), w \leftarrow f(a), z \leftarrow f(a)\}$$

questo può essere ottenuto applicando l'mgu seguito dall'ulteriore sostituzione

$$\{w \leftarrow f(a)\}$$

Proposizione 10.3. *L'algoritmo termina e produce, se esiste, l'unificatore più generale.*

Dimostrazione. (cenno) La terminazione dell'algoritmo è garantita dal fatto che le nuove equazioni che si generano contengono sempre meno simboli funzionali, quindi il passo 3. può essere iterato soltanto un numero finito di volte. Il passo 4. porta all'eliminazione di una variabile. I passi 1. e 2. possono essere usati al massimo per il numero di volte corrispondenti alle variabili. Con un semplice ragionamento induttivo si può concludere.

Riguardo la correttezza dell'algoritmo,

osserviamo innanzitutto che i due casi nei quali l'algoritmo dà fallimento sono evidentemente casi nei quali l'unificatore non può esistere. Inoltre, i passi dell'algoritmo trasformano il sistema di equazioni in un altro sistema che ha esattamente la stessa “soluzione”; ciò è banale per i primi tre passi ed è facilmente verificabile per il quarto, perché non fa altro che sostituire variabili con termini forniti dallo stesso sistema di equazioni, fino a fornire la soluzione in forma esplicita. Infine, l'unificatore così ottenuto è nella forma più generale perché non si è fatto uso di alcuna istanziazione che non fosse già implicita nel sistema di equazioni dato.

10.2.3 Risoluzione

Passo di risoluzione generale. Siano C e D due clausole senza variabili in comune tali

che $l_1 \in C$ e $l_2 \in D$ tali che l_1 ed l_2^c siano unificabili; allora diremo *risolvente* di C e D la clausola $\text{Res}[C, D]$ contenente tutti i letterali di $C\sigma$ tranne $l_1\sigma$ e tutti i letterali di $D\sigma$ tranne $l_2\sigma$, dove σ indica l'unificatore più generale.

Lemma 10.1. *La risolvente $\text{Res}[C, D]$ è soddisfacibile se lo sono le genitrici C e D .*

Dimostrazione. La dimostrazione richiede strumenti che non abbiamo introdotto per motivi di semplicità.

Procedura risolutiva generale. Sia S una data formula in forma clausale, cioè un insieme di clausole. Dovremo costruire una sequenza $S_0 \dots S_n$ di formule in forma clausale, partendo da $S_0 = S$ e costruendo ripetutamente la risolvente di due clausole contrastanti.

Consideriamo la formula S_i ; scegliamo due clausole contrastanti C e D di S_i e consideriamo la risolvente $\text{Res}[C, D]$. Costruiamo allora:

$$S_{i+1} = S_i \cup \text{Res}[C, D]$$

Se $\text{Res}[C, D]$ è la clausola vuota, $[]$, allora la procedura termina e S è insoddisfacibile. Se $S_{i+1} = S_i$ per ogni scelta di clausole contrastanti, allora la procedura termina e S è soddisfacibile. Altrimenti si itera con la costruzione di S_{i+2} .

Teorema 10.5 (Correttezza della risoluzione). *Se dalla formula in forma clausale S viene derivata mediante la procedura di risoluzione la clausola insoddisfacibile $[]$, allora S è insoddisfacibile.*

Dimostrazione. La dimostrazione si fa nel modo consueto: per induzione sul numero dei passi di risoluzione, usando il Lemma 10.1.

Teorema 10.6 (Completezza della risoluzione). *Se la formula in forma clausale S è insoddisfacibile, allora la clausola insoddisfacibile [] viene derivata mediante la procedura di risoluzione.*

La dimostrazione del teorema di completezza non è affatto semplice e richiede strumenti che non abbiamo introdotto in questo testo. Va però notato che il metodo presentato non costituisce una procedura di decisione, perché, se l'insieme di clausole è soddisfacibile, potrebbe non terminare. Questo è dovuto, come al solito, alla possibilità di modelli infiniti, come illustrato dal seguente controesempio.

Esempio 10.11. Consideriamo la formalizzazione del principio di induzione sui naturali:

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

Messa in forma clausale, tale formula diviene:

$$S_0 = \{ P(0), \neg P(x) \vee P(s(x)) \}$$

Applicando l'unico passo di risoluzione possibile, otteniamo

$$S_1 = \{ P(0), \neg P(x) \vee P(s(x)), P(s(0)) \}$$

Applicando l'unico passo di risoluzione che incrementa l'insieme di clausole, otteniamo

$$S_2 = \{ P(0), \neg P(x) \vee P(s(x)), P(s(0)), P(s(s(0))) \}$$

e così via, senza mai terminare.

Inoltre, per aumentare l'efficienza della risoluzione si ricorre spesso a strategie che impongono limitazioni all'applicazione del metodo. Non entriamo qui nei diversi casi, ma si tenga presente che in questo modo può diminuire la capacità dimostrativa del metodo; quindi la correttezza sarà comunque conservata, ma ogni volta si dovrà dimostrare un teorema di completezza. Vedremo nelle prossime due sezioni alcuni esempi di strategie di risoluzione. Forniamo intanto un esempio chiarificatore.

Esempio 10.12. Si consideri la seguente formula ottenuta dalla congiunzione di:

$$\begin{aligned} & \exists x \exists y P(x, y) \\ & \forall x \forall y \forall z ((P(x, y) \\ & \quad \wedge P(y, z)) \rightarrow P(x, z)) \end{aligned}$$

$$\begin{aligned} \forall x \forall y (P(x, y) \rightarrow P \\ (y, x)) \end{aligned}$$

Messa in forma clausale generale:

$$\begin{aligned} \{P(a, b), \neg P(x, y) \vee \neg P(y, z) \vee P(x, z), \\ \neg P(x, y) \vee P(y, x)\} \end{aligned}$$

Non è difficile dimostrare che $P(b, a)$ è conseguenza logica di S : basta risolvere la negazione di ciò che si vuole dimostrare e $\neg P(x, y) \vee P(y, x)$ tramite il mgu $\{x \leftarrow a, y \leftarrow b\}$ per ottenere $\neg P(a, b)$ il quale, risolto con $P(a, b)$, genera [].

Supponiamo però di adottare la strategia di risolvere sempre la prima coppia possibile di clausole e di letterali all'interno di tali clausole; diciamo inoltre di mettere all'inizio dell' $(i + 1)$ -esima forma clausale la risolvente ottenuta al passo i -esimo. Nel nostro esempio, partendo da

$$S_0 = \{\neg P(b, a), P(a, b), \neg P(x, y) \vee \neg P(y, z) \vee P(x, z), \neg P(x, y) \vee P(y, x)\}$$

ciò porterebbe a risolvere $\neg P(b, a)$ con $\neg P(x, y) \vee \neg P(y, z) \vee P(x, z)$ tramite il mgu $\{x \leftarrow b, z \leftarrow a\}$ e ottenere:

$$S_1 = \{\neg P(b, y) \vee \neg P(y, a)\} \cup S_0$$

dove, per poter procedere con maggior chiarezza, abbiamo ridenominato la y nella nuova clausola con una y nuova.

Ora dobbiamo risolvere la nuova clausola con $\neg P(x, y) \vee \neg P(y, z) \vee P(x, z)$ tramite il mgu $\{x \leftarrow b, z \leftarrow y'\}$; otteniamo così:

$$S_2 = \{\neg P(y', a) \vee \neg P(b, y'') \vee \neg P(y'', y')\} \cup S_1$$

con y'' nuova. Dobbiamo ancora risolvere la nuova clausola con $\neg P(x, y) \vee \neg P(y,$

$\exists z \vee P(x, z)$ tramite il mgu $\{x \leftarrow y', z \leftarrow a\}$; otteniamo così:

$$S_3 = \{\neg P(b, y'') \vee \neg P(y'', y') \vee \neg P(y', y''') \vee \neg P(y''', a)\} \cup S_2$$

con y''' nuova. E così via.

10.2.4 Dimostrazione automatica

Come abbiamo visto, neanche la procedura di risoluzione risolve sostanzialmente i nostri problemi di decidibilità e nemmeno, in generale, quelli di efficienza. Vedremo, però, in questo paragrafo, come permetta di affrontare il problema della dimostrazione automatica, cioè affidata ad una macchina.

Abbiamo visto un esempio di teoria assiomatica alla fine del precedente capitolo: quella dell'aritmetica di Peano.

Ma possiamo considerare teorie molto più semplici e, comunque, di interesse per il nostro scopo di illustrare la derivazione automatica dei teoremi. Si voglia, ad esempio, definire una teoria per le *successioni* di simboli di un certo alfabeto. Il linguaggio dovrà essere così fatto:

- il simbolo di costante ϵ per la successione vuota;
- un insieme finito di simboli di costanti a, b, c, \dots per i caratteri dell'alfabeto;
- il simbolo di funzione binario per la funzione di concatenazione (in pratica, non useremo mai esplicitamente tale simbolo e scriveremo la concatenazione di successioni semplicemente giustapponendole);
- il simbolo predicativo binario \sqsubseteq per la relazione di sottosuccessione.

Un insieme di assiomi, certamente non

sufficiente a provare tutte le proprietà che ci appetiamo come teoremi, è il seguente:

$$A_1 x = \epsilon x$$

$$A_2 x = y \rightarrow (zx = zy)$$

$$A_3 x = y \rightarrow (xz = yz)$$

$$A_4 x \sqsubseteq x$$

$$A_5 ((y = z_1 z_2) \wedge (x \sqsubseteq z_1)) \rightarrow (x \sqsubseteq y)$$

$$A_6 ((y = z_1 z_2) \wedge (x \sqsubseteq z_2)) \rightarrow (x \sqsubseteq y)$$

Supponiamo di voler dimostrare che la formula

$$G = abc \sqsubseteq aabcc$$

è un teorema della teoria delle successioni. Dobbiamo mostrare che G è una conseguenza logica degli assiomi:

$$A_1, \dots, A_6 \models G$$

Questo è vero se

$$A_1 \wedge \cdots \wedge A_6 \rightarrow G$$

è valida, ossia se possiamo refutare la sua negazione

$$\neg(A_1 \wedge \cdots \wedge A_6 \rightarrow G)$$

che equivale a

$$(A_1 \wedge \cdots \wedge A_6 \wedge \neg G)$$

Ma gli assiomi possono essere scritti facilmente in forma clausale. Per esempio l'assioma 5 è

$$\neg(y = z_1 z_2) \vee \neg(x \sqsubseteq z_1) \vee (x \sqsubseteq y)$$

Dunque per dimostrare una formula è sufficiente: considerare la sua negazione, aggiungere la sua forma clausale alle

clausole degli assiomi ed eseguire la procedura di risoluzione fino a quando non si ottiene (speriamo) la clausola vuota. Nel caso della formula G precedente, si ottiene la refutazione che segue dove è lasciato al lettore il compito di verificare nei dettagli le sostituzioni:

- 7. $\neg(abc \sqsubseteq aabcc)$
- 8. $\neg(aabcc = z'_1 z'_2) \vee \neg(abc \sqsubseteq z'_2) \quad 7, A_6$
- 9. $\neg(abcc = z'_2) \vee \neg(abc \sqsubseteq z'_2) \quad 8, A_2$
- 10. $\neg(abc \sqsubseteq abcc) \quad 9, A_1$
- 11. $\neg(abcc = z''_1 z''_2) \vee \neg(abc \sqsubseteq z''_1) \quad 10, A_5$
- 12. $\neg(abc = z''_1) \vee \neg(abc \sqsubseteq z''_1) \quad 11, A_3$
- 13. $\neg(abc \sqsubseteq abc) \quad 12, A_1$
- 14. $[] \quad 13, A_4$

Supponiamo ora di voler mostrare che una formula più complessa

$$\exists w(w \sqsubseteq aabcc)$$

sia conseguenza degli assiomi. In termini di risoluzione, cerchiamo di refutare la negazione

$$\neg(A_1 \wedge \dots \wedge A_6 \rightarrow \exists w(w \sqsubseteq aabcc))$$

che si trasforma in

$$A_1 \wedge \dots \wedge A_6 \wedge \forall w \neg(w \sqsubseteq aabcc)$$

e dunque, nuovamente, abbiamo un insieme di clausole. Se la refutazione della risoluzione è eseguita come nel paragrafo precedente la clausola finale è

$$\neg(w \sqsubseteq abc)$$

Il singolo letterale nella clausola contrasta col singolo letterale dell'assioma 4, *purchè* sia eseguita la sostituzione

$$\{w \leftarrow abc\}$$

Non solo abbiamo dimostrato che $\exists w(w \sqsubseteq aabcc)$ è una conseguenza logica degli assiomi, ma abbiamo anche *calcolato* un

valore per w (nel nostro caso abc) tale che $w \sqsubseteq aabcc$ sia vero.

Gli esempi considerati sono molto semplici e si può facilmente immaginare come le cose possano complicarsi in casi più generali generando un numero veramente grande di clausole. Si è pensato allora di introdurre restrizioni all'uso della regola di risoluzione in modo da aumentarne l'efficienza, mantenendo, ovviamente la correttezza del metodo, ma sacrificandone, se necessario, la completezza.

Una prima considerazione è che possiamo evitare di applicare la refutazione tra due clausole che sappiamo già essere mutuamente soddisfacibili, ad esempio gli assiomi della teoria (che, speriamo, non siano in contraddizione tra loro).

Definizione 10.3 (Risoluzione con insieme di supporto). Sia S un insieme di clausole e sia T un sottoinsieme di S tale che $S \setminus T$ sia soddisfacibile. La regola di risoluzione non può essere applicata tra due clausole in $S \setminus T$.

Teorema 10.7. La risoluzione con insieme di supporto è completa.

Dimostrazione. Vedere [Chang e Lee 73] o [Loveland 78].

Definizione 10.4 (Risoluzione lineare). Chiamiamo clausola centrale la formula originale da refutare. In ogni passo della risoluzione lineare una delle clausole genitrici è la clausola centrale corrente e la risolvente diviene la nuova clausola centrale.

Definizione 10.5 (Risoluzione sull'ingresso). Una delle clausole

genitrici è sempre una delle clausole originali.

La risoluzione lineare è completa mentre non lo è la risoluzione sull'ingresso. Nel prossimo paragrafo vedremo che se si restringe la forma dell'insieme di clausole, allora la risoluzione sull'ingresso e quella lineare sono entrambe complete ed efficienti.

10.2.5 Programmazione logica

Vediamo ora di descrivere un procedimento standard per dimostrare con efficienza il fatto che una formula sia conseguenza logica di un insieme di assiomi. In questo paragrafo, per uniformarci alla consuetudine della programmazione logica useremo l'implicazione inversa “ \Leftarrow ” invece dell'implicazione “ \rightarrow ”, più comune in

logica.

Osserviamo che di solito gli assiomi di una teoria si presentano come implicazioni con più (magari nessuna) premesse ed una sola conseguenza, cioè:

$$\forall x_1 \dots \forall x_k (B_1 \wedge \dots \wedge B_m \rightarrow B)$$

In forma clausale, un assioma è dunque:

$$\neg B_1 \vee \dots \vee \neg B_m \vee B$$

Per dimostrare che una formula G è una conseguenza logica di un insieme di assiomi aggiungiamo $\neg G$ all'insieme di assiomi e cerchiamo di costruire una refutazione mediante risoluzione. $\neg G$ è chiamata *clausola goal*. Come abbiamo detto alla fine del paragrafo precedente, non ha senso tentare di risolvere gli assiomi fra loro, così risolviamo la clausola goal con un assioma. Guardando

la forma degli assiomi, poichè $\neg G$ è un letterale negativo, esso si può risolvere solamente con un *singolo* letterale dell'assioma ed il risultato è una clausola i cui letterali sono negativi

$$\neg B_1 \vee \dots \vee \neg B_m$$

Così tutte le risolventi conterranno soltanto dei letterali negativi e non potranno essere risolte fra loro. Dato che è necessario fornire soltanto una risolvente in ogni momento, l'implementazione di questa forma di risoluzione è estremamente efficiente.

Descriviamo ora più formalmente questo metodo come un vero metodo di programmazione.

Definizione 10.6. Una clausola di Horn è una clausola con al massimo un letterale positivo:

$$A \leftarrow B_1, \dots, B_n$$

dove il letterale positivo A è chiamato la testa e quelli negativi B_i formano il corpo.

Ad una clausola di Horn positiva ed unitaria A si dà il nome di fatto. Una clausola di Horn senza letterali positivi $\leftarrow B_1, \dots, B_n$ è detta goal.

Definizione 10.7. *Un insieme di clausole di Horn non goal le cui teste abbiano la stessa lettera predicativa è chiamato procedura. Un programma è un insieme di procedure. Qualche volta una procedura composta soltanto da fatti chiusi è chiamata base di dati.*

Un esempio è mostrato nella [Figura 10.1](#). Questo programma ha due procedure (una per la testa $q(-, -)$ e l'altra per la testa $p(-, -)$), di cui una è una base di dati (la seconda).

| | | | |
|----|---------------------------------------|-----|-----------|
| 1. | $q(x, y) \Leftarrow p(x, y)$ | | |
| 2. | $q(x, y) \Leftarrow p(x, z), q(z, y)$ | | |
| 3. | $p(b, a)$ | 7. | $p(f, b)$ |
| 4. | $p(c, a)$ | 8. | $p(h, g)$ |
| 5. | $p(d, b)$ | 9. | $p(i, h)$ |
| 6. | $p(e, b)$ | 10. | $p(j, h)$ |

Figura 10.1 Un programma logico

Definizione 10.8. Sia P un programma e G un goal. Una sostituzione θ per le variabili in G è detta sostituzione per risposta corretta se $P \models \forall (\neg G \theta)$, dove “ \forall ” denota la chiusura universale delle variabili libere in $\neg G \theta$.

Per esempio, se P è l’insieme degli assiomi dell’aritmetica e G è

$$\neg(6 + y = 13)$$

allora

$$\theta = \{y \leftarrow 7\}$$

è una sostituzione per risposta corretta perché

$$P \vDash (6 + 7 = 13)$$

Una sostituzione per risposta corretta non è necessariamente una sostituzione chiusa. Se ora G è

$$\neg(x = y + 13)$$

allora

$$\theta = \{y \leftarrow x - 13\}$$

è una sostituzione per risposta corretta in quanto

$$P \vDash \forall x (x = x - 13 + 13)$$

Osserviamo che la sostituzione θ è

necessaria dato che la chiusura di $\neg G$ non è una conseguenza logica di P :

$$P \not\models \forall x \ \forall y (x = y + 13)$$

In generale, supponiamo di volere scoprire se

$$B = \exists (B_1 \wedge \dots \wedge B_n)$$

è una conseguenza logica di un programma (insieme di clausole) P , dove \exists denota la chiusura esistenziale della congiunzione. Allora

$$P \models B$$

se e soltanto se esiste qualche sostituzione chiusa σ tale che

$$P \models (B_1 \wedge \dots \wedge B_n)\sigma$$

Supponiamo che σ possa essere scritta

come una composizione di sostituzioni $\sigma = \theta \lambda$, per *una qualche* sostituzione chiusa λ (in pratica, esiste sempre una tale composizione, nel peggior dei casi basta considerare $\theta = \sigma$ e $\lambda = e$). Segue che

$$P \vDash \forall ((B_1 \wedge \cdots \wedge B_n)\theta)$$

Così ‘ B è una conseguenza logica di P ’ è equivalente al fatto che θ sia una sostituzione per risposta corretta per la clausola goal

$$G = \neg(B_1 \wedge \cdots \wedge B_n)$$

e, poichè G è implicitamente quantificata universalmente,

$$\begin{aligned} \forall G &= \forall \neg(B_1 \wedge \cdots \wedge B_n) \\ &\equiv \neg \exists (B_1 \wedge \cdots \wedge B_n) \\ &= \neg B \end{aligned}$$

Dunque $P \models B$ se e soltanto se $\models P \rightarrow B$, ossia $\not\models \neg(P \rightarrow B)$; ciò equivale a $\not\models P \wedge \neg B$, cioè $\not\models P \wedge G$. Concludiamo che tutto ciò che è necessario per ottenere una risposta dal programma consiste nell'aggiungere come clausola goal la negazione della domanda e quindi usare la risoluzione per refutare l'insieme di clausole ottenuto. La sostituzione prodotta dall'unificazione definisce una sostituzione per risposta corretta.

Pertanto, gli assiomi di una teoria formano un programma in grado di calcolare risposte a domande. Questo programma è estremamente non deterministico:

- Data una clausola goal, possiamo decidere di risolvere uno dei letterali presenti nella clausola.
- Una volta che abbiamo scelto un

letterale, possiamo decidere di risolverlo con un qualsiasi assioma il cui letterale positivo contrasta con il letterale scelto.

Il formalismo non deterministico diventa un pratico linguaggio di programmazione logica specificando le regole che giustificano queste scelte.

Definizione 10.9. *Una regola di calcolo è una regola per scegliere i letterali in una clausola goal. Una regola di ricerca è una regola che sceglie una clausola dalla procedura per risolverla con il letterale scelto in una clausola goal.*

Definizione 10.10 (SLD-risoluzione). *Sia P un insieme di clausole programma, R una regola di calcolo, R' una regola di ricerca e G un goal. Una derivazione mediante SLD-risoluzione è definita come una*

successione di passi di risoluzione tra la clausola goal e le clausole programma:

- *la prima clausola goal G_0 è G .*
- *assumiamo di aver derivato G_i , ossia $\Leftarrow A_1, \dots, A_{i-1}, A_i, A_{i+1}, \dots, A_n$. Allora, G_{i+1} è $\Leftarrow (A_1, \dots, A_{i-1}, B_1, \dots, B_k, A_{i+1}, \dots, A_n) \theta_i$, dove il letterale A_i è selezionato secondo R , la clausola $A \Leftarrow B_1, \dots, B_k$ è selezionata secondo R' e θ_i è l'mgu tra A e A_i .*

Una SLD-refutazione è una SLD-derivazione di $[]$.

Teorema 10.8 (Correttezza della SLD-risoluzione). *Sia P un insieme di clausole-programma, R una regola di calcolo, R' una regola di ricerca e G un goal. Supponiamo che esista una SLD-refutazione di G . Sia*

$$\theta = \theta_1 \dots \theta_n$$

la composizione degli unificatori usati nella refutazione e sia σ la restrizione di θ alle variabili di G . Allora σ è una sostituzione per risposta corretta per G .

Dimostrazione. Per definizione di σ , $G\theta = G\sigma$, dunque $P \cup \{G\sigma\} = P \cup \{G\theta\}$ che è insoddisfacibile a causa della correttezza della risoluzione. Quindi $P \vDash \neg G\sigma$. Poichè ciò è vero per una qualsiasi sostituzione delle variabili libere di $G\sigma$, $P \vDash \forall(\neg G\sigma)$.

Osserviamo che la SLD-refutazione *non* è completa per insiemi di clausole qualsiasi. L'insieme di clausole

$$p \vee q, \neg p \vee q, p \vee \neg q, \neg p \vee \neg q$$

è insoddisfacibile ma non ha nessuna SLD-risoluzione. Tale procedimento,

infatti, impone che il goal del passo $(i + 1)$ -esimo sia ottenuto da quello del passo i -esimo. Invece, nell'esempio in questione, l'unico modo di risolvere l'insieme di clausole dato è il seguente:

1. risolviamo pq e \underline{pq} , ottenendo q ;
2. risolviamo pq e \underline{pq} , ottenendo q ;
3. risolviamo q e q , ottenendo [].

Si può notare che il secondo passaggio non rispetta i vincoli della SLD-risoluzione. Per le clausole di Horn, comunque, la SLD-risoluzione è completa.

Teorema 10.9 (Completezza della SLD-risoluzione). *Sia P un insieme di clausole-programma di Horn, R una regola di calcolo, R' una regola di ricerca e G un goal. Sia σ una sostituzione per risposta corretta. Allora c'è una*

SLD-refutazione di G da P tale che σ è la

restrizione della composizione degli unificatori

$$\theta = \theta_i \cdots \theta_n$$

alle variabili in G.

Dimostrazione. La dimostrazione di completezza, nonostante non sia difficile da seguire, richiede delle tecniche che vanno oltre l'ambito di questo libro. Si può vedere [Lloyd 84] per i dettagli.

Diamo ora un esempio usando il programma della [Figura 10.1](#) e la seguente clausola goal

$$\Leftarrow q(u, b), q(b, v)$$

Ad ogni passo possiamo scegliere un letterale della clausola e una clausola la cui testa contrasti con il letterale.

1. Scegliamo $q(u, b)$ e risolviamolo con la clausola 1 mediante la mgu $\{x \leftarrow u, y \leftarrow b\}$:

$$\Leftarrow p(u, b), q(b, v)$$

2. Scegliamo $p(u, b)$ e risolviamolo con la clausola 5. Ciò richiede la sostituzione $\{y \leftarrow d\}$:

$$\Leftarrow q(b, v)$$

3. Ora possiamo scegliere soltanto un letterale e risolverlo, per esempio, con la clausola 1 mediante la mgu $\{x \leftarrow b, y \leftarrow v\}$:

$$\Leftarrow p(b, v)$$

4. Possiamo scegliere soltanto un letterale e risolverlo, per esempio, con la clausola 3. Questo richiede la sostituzione $\{z \leftarrow a\}$ e genera $[]$.

Poichè la clausola goal è

$$\neg q(u, b) \vee \neg q(b, v)$$

segue che una sostituzione per risposta corretta è

$$\{u \leftarrow d, v \leftarrow a\}$$

e che

$$P \vDash q(d, b) \wedge q(b, a)$$

Dal calcolo dei predicati elementare segue che

$$P \vDash \exists u \exists v (q(u, b) \wedge q(b, v))$$

Per concludere, notiamo che la differenza tra programmazione logica e (ordinaria) programmazione algoritmica è che nella programmazione algoritmica il controllo del calcolo è costruito *esplicitamente* dal

programmatore come parte del programma stesso. Ciò risulta evidente dal fatto che le strutture di controllo, in linguaggi come il C, occupano un ruolo centrale. Nella programmazione logica, il programmatore scrive delle formule logiche dichiarative che descrivono la relazione tra l'ingresso e l'uscita e quindi il ‘compilatore’, cioè il motore di inferenza della risoluzione insieme con le regole di ricerca e di calcolo, fornisce una struttura di controllo *uniforme*.

Ovviamente, per quanto efficiente sia un compilatore di programmazione logica, la sua struttura uniforme di controllo non potrà mai competere con una struttura di controllo ‘fatta a mano’ per un calcolo specifico. La ricerca in programmazione logica ha esplorato i costi e i benefici dell’uso di logica dichiarativa ‘pura’ in contrasto con i compromessi procedurali

‘impuri’ che permettono al programmatore di scrivere programmi la cui efficienza è pari a quella dei linguaggi algoritmici.

Esercizi svolti

Esercizio 10.1. *Sia S una forma clausale e l un letterale tale che l appare in S e \bar{l} non appare in S ; si dimostri che, detta T la formula in forma clausale ottenuta da S cancellando ogni clausola che contiene l , risulta $S \approx T$.*

Soluzione: Sia $S = \{C_1, \dots, C_k\}$ e $T = \{C_{j_1}, \dots, C_{j_h}\}$, per $\{j_1, \dots, j_h\} \subseteq \{1, \dots, k\}$ e C_n che contiene l per ogni $n \in \{1, \dots, k\} \setminus \{j_1, \dots, j_h\}$. Se S è soddisfacibile, allora, banalmente, lo stesso assegnamento di verità soddisfa T , visto che $T \subseteq S$ e,

quindi, ogni clausola di T ha almeno un letterale che vale “vero”. Viceversa, se T è soddisfacibile, allora basta estendere l’assegnamento di verità che soddisfa T in modo da assegnare valore “vero” a l ; si ottiene così un assegnamento che soddisfa ogni clausola di S .

Esercizio 10.2. *Sia S una forma clausale e $\{l\}$ una clausola unitaria di S ; si dimostri che, detta T la formula in forma clausale ottenuta da S cancellando ogni clausola che contiene l e inoltre cancellando l^c all’interno di ogni clausola rimanente, risulta $S \approx T$.*

Soluzione: Sia $S = \{\{l\}, C_1, \dots, C_k\}$ e $T = \{C'_{j_1}, \dots, C'_{j_h}\}$, per $\{j_1, \dots, j_h\} \subseteq \{1, \dots, k\}$, C_n che contiene l per ogni $n \in \{1, \dots, k\} \setminus \{j_1, \dots, j_h\}$ e C'_{j_m} ottenuta da C_{jm} rimuovendo l^c per ogni $m \in \{1, \dots, h\}$. Se S è soddisfacibile,

allora ogni assegnamento che la soddisfa deve assegnare valore “vero” a l (per via della clausola unitaria $\{l\}$ in S), e quindi “falso” a l^c ; pertanto, in tutte le clausole di T c’è almeno un letterale che vale “vero”, quindi T è soddisfacibile. Viceversa, se T è soddisfacibile, per ottenere un assegnamento che soddisfi S basta estendere un assegnamento che soddisfa T ponendo l a “vero”: questo soddisferà ogni clausola che contiene l , mentre i restanti assegnamenti soddisferanno tutte le altre clausole (visto che già le soddisfacevano in T).

Esercizio 10.3. *Data la forma clausale $S_0 = \{\underline{P} \ Q \ R, P \ R, Q \ R, \underline{R}\}$, se ne costruiscano due diverse refutazioni.*

Soluzione: Iniziamo col risolvere prima $\underline{P} \ Q \ R$ e $P \ R$ rispetto a P : ciò genera $S_1 = S_0 \cup \{Q \ R\}$. Risolviamo ora $Q \ R$ e $\underline{Q} \ R$

rispetto a Q : ciò genera $S_2 = S_1 \cup \{R\}$.

Risolviamo ora R e \underline{R} rispetto a R : ciò genera $[]$ e permette di concludere.

Proviamo ora a iniziare risolvendo QR e \underline{R} rispetto a R : ciò genera $S'_1 = S_0 \cup \{Q\}$. Risolviamo ora Q e $\underline{P} \ Q \ R$ rispetto a Q : ciò genera $S'_2 = S'_1 \cup \{\underline{P} \ R\}$. Risolviamo ora $P \ R$ e $\underline{P} \ R$ rispetto a P : ciò genera $S'_3 = S'_2 \cup \{R\}$. Risolviamo infine R e \underline{R} rispetto a R : ciò genera $[]$ e permette di concludere.

Esercizio 10.4. *Si provi a dimostrare che $((P \rightarrow Q) \vee \neg Q) \rightarrow (P \rightarrow R)$ è soddisfacibile usando il metodo di risoluzione.*

Soluzione: Anzitutto, portiamo la formula data in forma clausale. L'applicazione dei cinque passi descritti in Sezione 10.1.1 al nostro esempio sono (in realtà, il quinto punto nel nostro caso non cambia la

formula):

1. $\neg(\neg P \vee Q \vee \neg Q) \vee (\neg P \vee R);$
2. $(\neg\neg P \wedge \neg Q \wedge \neg\neg Q) \vee (\neg P \vee R);$
3. $(P \wedge \neg Q \wedge Q) \vee (\neg P \vee R);$
4. $(P \vee \neg P \vee R) \wedge (\neg Q \vee \neg P \vee R) \wedge (Q \vee \neg P \vee R).$

Pertanto, la forma clausale cercata è $S = \{P \underline{P} R, \underline{P} Q R, \underline{P} Q R\}$. Se proviamo a risolvere $\underline{P} Q R$ e $\underline{P} Q R$ rispetto a Q , otteniamo $\underline{P} R$, da cui $S' = S \cup \{\underline{P} R\}$. Se ora proviamo a risolvere $\underline{P} P R$ e $\underline{P} Q R$ rispetto a P , otteniamo $\underline{P} Q R$, che è già presente in S ; se proviamo a risolvere $\underline{P} P R$ e $\underline{P} Q R$ rispetto a P , otteniamo $\underline{P} Q R$, che è già presente in S ; se proviamo a risolvere $\underline{P} P R$ e $\underline{P} R$ rispetto a P , otteniamo $\underline{P} R$, che è già presente in S ; se proviamo a risolvere $\underline{P} Q R$ e $\underline{P} Q R$ rispetto a Q , otteniamo $\underline{P} R$, che è già presente in S' . Siccome ogni possibile

passo di risoluzione non cambia la forma clausale S' , per il teorema 10.3 abbiamo che S è soddisfacibile. Infatti, se non lo fosse, dovremmo poter derivare $[]$, cosa che non siamo stati in grado di fare.

Esercizio 10.5. *Si dimostri, mediante risoluzione, che $\neg A \wedge \neg B$ è conseguenza logica di $(A \rightarrow \neg B) \wedge (\neg A \rightarrow \neg B) \wedge (A \rightarrow B)$.*

Soluzione: Anzitutto, portiamo la formula $(A \rightarrow \neg B) \wedge (\neg A \rightarrow \neg B) \wedge (A \rightarrow B) \wedge \neg(\neg A \wedge \neg B)$ in forma clausale; otteniamo $S_0 = \{\underline{A} \underline{B}, A \underline{B}, \underline{A} B, AB\}$. Risolviamo $\underline{A} \underline{B}$ e $A \underline{B}$ su A e otteniamo $S_1 = S_0 \cup \{\underline{B}\}$.

Risolviamo $\underline{A} B$ e AB su A e otteniamo $S_2 = S_1 \cup \{B\}$. Risolviamo B e \underline{B} su B e otteniamo $[]$.

Esercizio 10.6. *Si applichi l'algoritmo di unificazione alle formule $P(a, y, f(y))$ e P*

(z, z, u) e si trovi, se possibile, il loro mgu.

Soluzione: All'inizio, si considera la sola equazione

$$P(a, y, f(y)) = P(z, z, u)$$

Si applica il passo 3 dell'algoritmo e si ottiene così il nuovo sistema di equazioni

$$a = z$$

$$y = z$$

$$f(y) = u$$

Applichiamo il passo 1 per ottenere

$$z = a \quad y = z \quad u = f(y)$$

Applichiamo il passo 4 alla prima equazione per ottenere

$$z = a \quad y = a \quad u = f(y)$$

e riapplichiamo il passo 4 alla seconda equazione per concludere

$$z = a \quad y = a \quad u = f(a)$$

Infatti, la sostituzione $\{z \leftarrow a, y \leftarrow a, u \leftarrow f(a)\}$ unifica le due formule, producendo $P(a, a, f(a))$.

Esercizio 10.7. *Si applichi l'algoritmo di unificazione alle formule $P(x, f(y))$ e $P(g(y), x)$ e si trovi, se possibile, il loro mgu.*

Soluzione: All'inizio, si considera la sola equazione

$$P(x, f(y)) = P(g(y), x)$$

Si applica il passo 3 dell'algoritmo e si ottiene così il nuovo sistema di equazioni

$$x = g(y)$$

$$f(y) = x$$

Applichiamo il passo 1 per ottenere

$$x = g(y)$$

$$x = f(y)$$

Applichiamo il passo 4 alla prima equazione per ottenere

$$x = g(y)$$

$$g(y) = f(y)$$

A questo punto, l'algoritmo termina dichiarando che le formule date non sono unificabili, poichè $f \# g$.

Esercizio 10.8. *Si dimostri, usando il metodo di risoluzione, che*

$$\forall x (\exists y (P(x) \rightarrow Q(y)) \wedge \exists y \neg(P(x) \rightarrow Q(y)))$$

è insoddisfacibile.

Soluzione: Dobbiamo anzitutto portare la formula in forma clausale.

Ridenominiamo la y del terzo quantificatore in z ; eliminiamo poi le

implicazioni, spingiamo dentro le negazioni e tiriamo fuori tutti i quantificatori. Otteniamo così

$$\forall x \exists y \exists z ((\neg P(x) \vee Q(y)) \wedge P(x) \wedge \neg Q(z))$$

che è in forma normale congiuntiva prenessa. Skolemizzando tale formula, otteniamo la forma clausale

$$S_0 = (\neg P(x) \vee Q(f(x))), P(x), \neg Q(f(x))$$

Applichiamo la risoluzione e risolviamo dapprima $\neg P(x) \vee Q(f(x))$ e $P(x)$; otteniamo così

$$S_1 = S_0 \cup \{Q(f(x))\}$$

La risolvente di $Q(f(x))$ e $\neg Q(f(x))$ genera [] e ci permette di concludere.

Esercizio 10.9. Consideriamo il paradosso del barbiere già descritto nella Nota 4.1. Assumiamo che

1. ogni barbiere rade chi non si rade da sè;
2. nessun barbiere rade chi si rade da sè.

Si mostri come queste due affermazioni implicano che non esiste nessun barbiere.

Soluzione: Definiamo $B(x)$ il predicato “ x è un barbiere” e $R(x, y)$ il predicato “ x rade y ”. Le tre affermazioni in esame vengono quindi formalizzate come segue:

1. $\forall x \forall y((B(x) \wedge \neg R(y, y)) \rightarrow R(x, y));$
2. $\neg \exists x \exists y(B(x) \wedge R(y, y) \wedge R(x, y));$
3. $\neg \exists x B(x).$

Poichè siamo interessati a mostrare che la terza è conseguenza logica delle prime

due, dobbiamo considerare la sua negata:

$$3' \exists x B(x).$$

Mettiamo quindi (1), (2) e (3') in forma clausale, ottenendo

$$S_0 = \{(\neg B(x) \vee R(y, y) \vee R(x, y)), (\neg B(x) \vee \neg R(y, y) \vee \neg R(x, y)), B(a)\}$$

dove a è una funzione di arità 0 (cioè, una costante) nuova.

Risolviamo ora le prime due clausole, ridenominando le variabili in (2) in modo da renderle diverse da quelle in (1); useremo x' invece di x e y' invece di y . Si tratta quindi di unificare $R(x, y)$ con $\neg R(y', y')$: ciò genera il mgu $\{x \leftarrow y', y \leftarrow y'\}$ che porta a

$$S_1 = S_0 \cup \{\neg B(x') \vee \neg B(y') \vee \neg R(x', y')\}$$

Risolviamo ora il letterale $\neg R(x', y')$ della nuova clausola con $R(y, y)$ di (1); otteniamo il mgu $\{x' \leftarrow y, y' \leftarrow y\}$ e

$$S_2 = S_1 \cup \{\neg B(y) \vee \neg B(x) \vee R(x, y)\}$$

Risolviamo ora le due risolventi appena ottenute sui letterali $\neg R(x', y')$ e $R(x, y)$; otteniamo il mgu $\{x' \leftarrow x, y' \leftarrow y\}$ e

$$S_3 = S_2 \cup \{\neg B(y) \vee \neg B(x)\}$$

Risolviamo ora l'ultima clausola con $B(a)$; otteniamo il mgu $\{y \leftarrow a\}$ e

$$S_4 = S_3 \cup \{\neg B(x)\}$$

Risolviamo ora l'ultima clausola con $B(a)$; otteniamo il mgu $\{x \leftarrow a\}$ e $[]$, come richiesto.

Esercizio 10.10. Consideriamo i

seguenti fatti:

1. *tutti gli studenti che si presentano all'esame di logica sono interrogati da un professore di logica;*
2. *alcuni studenti sono appassionati di logica e sono interrogati soltanto da appassionati di logica;*
3. *tutti gli appassionati di logica si presentano all'esame di logica.*

Si mostri come queste affermazioni implicano che alcuni professori di logica sono appassionati di logica.

Soluzione: Definiamo: $S(x)$ il predicato “ x è uno studente”; $E(x)$ il predicato “ x si presenta all’esame di logica”; $I(x, y)$ il predicato “ x interroga y ”; $P(x)$ il predicato “ x è un professore di logica”; $A(x)$ il predicato “ x è un appassionato di logica”. Le quattro affermazioni in esame vengono

ora formalizzate come segue:

1. $\forall x ((S(x) \wedge E(x)) \rightarrow \exists y(P(y) \wedge I(y, x)))$;
2. $\exists x (S(x) \wedge A(x) \wedge \forall y(I(y, x) \rightarrow A(y)))$;
3. $\forall x (A(x) \rightarrow E(x))$.
4. $\exists x (P(x) \wedge A(x))$.

Poichè siamo interessati a mostrare che la quarta è conseguenza logica delle prime tre, dobbiamo considerare la sua negata:

4. $\neg \exists x (P(x) \wedge A(x))$.

Mettiamo quindi (1), (2), (3) e (4') in forma clausale, ottenendo

$$\begin{aligned} S_0 = \{ & \neg S(x) \vee \neg E(x) \vee P(f(x)), \\ & \neg S(x) \vee \neg E(x) \vee I(f(x), x), \\ & S(a), A(a), \neg I(y, a) \vee A(y), \\ & \neg A(x) \vee E(x), \neg P(x) \vee \neg A(x) \\ \} \end{aligned}$$

Da $S(a)$ e $\neg S(x) \vee \neg E(x) \vee P(f(x))$
otteniamo

$$S_1 = S_0 \cup \{\neg E(a) \vee P(f(a))\}$$

Da $S(a)$ e $\neg S(x) \vee \neg E(x) \vee I(f(x), x)$
otteniamo

$$S_2 = S_1 \cup \{\neg E(a) \vee I(f(a), a)\}$$

Da $A(a)$ e $\neg A(x) \vee E(x)$ otteniamo

$$S_3 = S_2 \cup \{E(a)\}$$

Da $E(a)$ e $\neg E(a) \vee P(f(a))$ otteniamo

$$S_4 = S_3 \cup \{P(f(a))\}$$

Da $E(a)$ e $\neg E(a) \vee I(f(a), a)$ otteniamo

$$S_5 = S_4 \cup \{I(f(a), a)\}$$

Da $I(f(a), a)$ e $\neg I(y, a) \vee A(y)$ otteniamo

$$S_6 = S_5 \cup \{A(f(a))\}$$

Da $A(f(a))$ e $\neg P(x) \vee \neg A(x)$ otteniamo

$$S_7 = S_6 \cup \{\neg P(f(a))\}$$

Da $P(f(a))$ e $\neg P(f(a))$ otteniamo [].

Esercizio 10.11. *Sia dato il seguente programma:*

$$\begin{aligned} &P(a, b) \\ &P(c, b) \\ &P(x, y) \Leftarrow P(x, z), P(z, y) \\ &P(x, y) \Leftarrow P(y, x) \end{aligned}$$

Si dimostri che da esso è possibile derivare il goal $P(a, c)$ e che tutte e quattro le clausole del programma sono

necessarie a questo scopo.

Soluzione: Unifichiamo il goal con la testa della terza clausola del programma, ottenendo così

$$\Leftarrow P(a, z), P(z, c)$$

Unifichiamo ora $P(a, z)$ con la prima clausola del programma e otteniamo

$$\Leftarrow P(b, c)$$

Unifichiamo $P(b, c)$ con la testa dell'ultima clausola del programma per ottenere

$$\Leftarrow P(c, b)$$

da cui, grazie alla seconda clausola del programma, otteniamo [].

È facile convincersi che, comunque si prenda un sottoprogramma proprio del

programma dato, il goal non è più derivabile. Ciò basta per concludere.

Esercizi da svolgere

Esercizio 10.12. *Sia S una forma clausale contenente una clausola C cui appartengono sia l che $\neg l$; si dimostri che, detta T la formula in forma clausale ottenuta da S cancellando C , risulta $S \approx T$.*

Esercizio 10.13. *Sia S una forma clausale e siano C e D due clausole di S tali che $C \subseteq D$; si dimostri che, detta T la formula in forma clausale ottenuta da S cancellando D , risulta $S \approx T$ (cioè in questo caso possiamo cancellare la clausola “più grande”).*

Esercizio 10.14. *Si provino, usando il metodo di risoluzione per il calcolo degli*

enunciati, le equivalenze logiche dell'Esempio 8.7.

Esercizio 10.15. *Si dimostri mediante risoluzione che $P \rightarrow R$ è conseguenza logica di $P \wedge (Q \rightarrow R) \wedge (P \rightarrow Q)$.*

Esercizio 10.16. *Si dimostri mediante risoluzione che A è conseguenza logica di $((\neg A \rightarrow B) \vee C) \wedge (C \rightarrow B) \wedge (B \rightarrow A)$.*

Esercizio 10.17. *Si provino ad unificare le seguenti coppie di formule e, in caso di successo, si dia il loro mgu:*

1. $P(x, y) \text{ e } P(y, f(z))$;
2. $Q(x, g(x)) \text{ e } Q(y, y)$;
3. $R(x, g(x), y) \text{ e } R(z, u, g(a))$;
4. $R(x, g(x), y) \text{ e } R(a, g(a), v)$;
5. $R(z, u, g(a)) \text{ e } R(a, g(a), v)$;
6. $P(a, x, f(g(y))) \text{ e } P(y, f(z), f(z))$;
7. $P(a, x, f(g(y))) \text{ e } P(z, h(z, u), f(u))$.

Esercizio 10.18. Si trasformi in forma clausale la formula

$$\forall x \exists y (\neg R(y, x) \vee \exists z \neg(Q(x, z) \rightarrow P(y, z)))$$

Esercizio 10.19. Assumiamo che ogni uomo è contento se ogni suo figlio è obbediente. Da ciò si dimostri che ogni uomo senza figli è contento.

Esercizio 10.20. Si assuma che ci sono cittadini che apprezzano tutti i governanti, ma che nessun cittadino apprezza i tiranni. Si mostri che una conseguenza logica di questi fatti è che non ci sono governanti tiranni.

Prospettive e approfondimenti

Il mondo della logica matematica è enormemente più vasto dell'introduzione che ne abbiamo presentato in questo testo. Qua e là abbiamo avvertito che stavamo facendo scelte semplificative importanti: ad esempio il numero dei valori di verità, gli operatori consentiti (connettivi e quantificatori), i sistemi deduttivi, le tecniche di risoluzione, ecc. Ognuna di queste scelte ci ha escluso un orizzonte a dir poco sterminato.

La logica matematica ha conosciuto nel secolo scorso uno sviluppo enorme, essendole stato richiesto di intervenire a portare rigore nei campi più disparati: matematica, filosofia, giurisprudenza, ecc. Ogni volta l'adattarsi ai diversi linguaggi ha provocato lo sviluppo di tecniche specifiche che l'hanno diversificata ed arricchita. Ad esempio, un uso diverso del legame tra i connettivi ed i valori di verità ha portato allo sviluppo della *logica intuizionista*, così importante nello studio dei fondamenti della matematica. In bibliografia il lettore interessato potrà trovare riferimenti per ampliare le sue conoscenze.

Dato lo scopo di questo volume, vale la pena accennare ad alcune delle “logiche” che hanno trovato fortuna ed applicazione nell’informatica. Dal momento che la logica matematica si interessa soprattutto

di linguaggio e di calcolo, è naturale che l'informatica le abbia offerto un ambiente propizio per utilizzare ed affinare le sue tecniche. Anzi, forse proprio dalle applicazioni all'informatica, è diventato evidente che a diversi linguaggi corrispondano naturalmente logiche diverse, anche se tutte rigorose da un punto di vista matematico. Questo è ormai un fatto scontato, ma non è stato sempre così: per tanti secoli il tentativo di usare una logica unica per diversi linguaggi ha condotto a paradossi ed antinomie.

Abbiamo parlato nel testo delle applicazioni nella dimostrazione automatica che hanno portato a sviluppare le tecniche di risoluzione. Da qui si possono facilmente intuire sviluppi nel campo dell'intelligenza artificiale. Non abbiamo invece accennato ad un'altra applicazione importante che sta avendo

successo in questo momento: la specifica formale e la verifica di sistemi informatici. Qui la logica permette in maniera rigorosa (ed economica) di evitare e correggere errori di progettazione. In questo ambito hanno trovato applicazione le *logiche modali*, cioè quelle che utilizzano operatori che non sono connettivi né quantificatori, ma sono capaci di esprimere modalità di validità/soddisfabilità di una formula in relazione ad un insieme che possiede una struttura. Si pensi ad una formula che esprima proprietà che dipendono dagli stati di un sistema: il valore della formula cambierà passando da uno stato all'altro, ma se questi stati sono legati da una relazione (ad esempio temporale), vorremmo in certi casi affermare magari che il suo valore resta invariato per tutti gli stati “futuri”, oppure che “ci sarà prima o poi” uno stato nel quale la formula

assumerà un certo valore. Si può trovare un'introduzione abbastanza elementare a questi argomenti in due testi di logica matematica per informatica molto più estesi del nostro [7, 25].

Esistono poi logiche più sofisticate dal punto di vista matematico, nate per motivi fondazionali, come quella *lineare* e quella *categoriale*, capaci tuttavia di esprimere e trattare anche in informatica concetti importanti come “risorsa” o “validità locale”. Per queste non diamo riferimenti bibliografici, perché sarebbero prematuri per i nostri lettori all'inizio del corso di studio. Speriamo che tra loro qualcuno si appassioni tanto a questi argomenti da arrivare fin là.

Bibliografia

- [1] Aigner, M. e Ziegler, G.M. (1998). *Proofs from The Book*. Springer, Berlin.
- [2] Aimonetto, I. (1975). *Le antinomie logiche e semantiche*. Filosofia, Torino.
- [3] Bagni, G.T. (1997). *Elementi di storia della logica formale*. Pitagora, Bologna.
- [4] Bagni, G.T. (2002). *Congetture e teorie aritmetiche*. Archimede, 2:96-100.
- [5] Bellacicco, A. e Labella, A. (1979). *Le strutture matematiche dei dati*.

Feltrinelli, Milano.

- [6] Ben-Ari, M. (1998). *Logica matematica per l'informatica*, a cura di A. Labella. UTET Università, Roma.
- [7] Ben-Ari, M. (2001). *Mathematical Logic for Computer Science*, 2nd edition. Springer Verlag.
- [8] Bochenski, J.M. (1972). *La logica formale, I-II* Einaudi, Torino.
- [9] Bottazzini, U. (1990). *Il flauto di Hilbert*. UTET, Torino.
- [10] Bourbaki, N. (1963). *Elementi di storia della matematica*. Feltrinelli, Milano.
- [11] Casari, E. (1964). *Questioni di filosofia della matematica*. Feltrinelli, Milano.
- [12] Casari, E. (1973). *La filosofia della matematica del '900*. Sansoni,

Firenze.

- [13] Casari, E. (a cura di) (1979). *Dalla logica alla metalogica*. Sansoni, Firenze.
- [14] Cellucci, C. (a cura di) (1967). *La filosofia della matematica*. Laterza, Bari.
- [15] Crossley, J.N. (1976). *Che cos'è la logica matematica?* Boringhieri, Torino.
- [16] D'Amore, B. e Matteuzzi, M. (1975). *Dal numero alla struttura*. Zanichelli, Bologna.
- [17] Falletta, N. (1989). *Il libro dei paradossi*. Longanesi, Milano.
- [18] Garciadiego, A.R. (1992). *Bertrand Russell and the Origins of the Set-theoretic “Paradoxes”*. Birkhäuser, Basel.
- [19] Gödel, K. (1979). *La completezza*

degli assiomi del calcolo logico funzionale. Dalla logica alla metalogica, a cura di E.Casari, 137-149, Sansoni, Firenze (ed. orig.: Monatshefte für Mathematik und Phisik, 38:349-360, 1930).

- [20] Gödel, K. (1931). *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme.* Monatshefte für Mathematik und Physik:38.
- [21] Goe, G. (1983). *Lezioni di logica.* Angeli, Milano.
- [22] Guy, R.K. (1994). *Unsolved Problems in Number Theory*, seconda edizione. Springer, New York.
- [23] Hamblin, H. (1970). *Fallacies.* Methuen, Londra.
- [24] Hilbert, D. (1979). *I fondamenti logici della matematica.* Dalla logica alla

metalogica, a cura di E.Casari, 67-78, Sansoni, Firenze (ed. orig.: *Mathematische Annalen*, 88:151-165, 1923).

- [25] Huth, M. e Ryan, M. (2004). *Logic in Computer Science*. Cambridge University Press.
- [26] Kline, M. (1991). *Storia del pensiero matematico, I-II* Einaudi, Torino.
- [27] Kneale, W.C. e Kneale, M. (1972). *Storia della logica*. Einaudi, Torino.
- [28] Koiré, A. (1947). *Epiménide le menteur*. Hermann, Parigi.
- [29] Labella, A. (2000). *Algorithmic and geometric thinking: the example of Campanus*. Categorical Studies in Italy. Supplemento ai Rendiconti del Circolo Matematico di Palermo II, 64:193-200.
- [30] Lolli, G. (1974). *Teoria assiomatica*

degli insiemi. Boringhieri, Torino.

- [31] Lolli, G. (1979). *Lezioni di logica matematica*. Boringhieri, Torino.
- [32] Lolli, G. (1985). *Le ragioni fisiche e le dimostrazioni matematiche*. Il Mulino, Bologna.
- [33] Lolli, G. (1988). *Introduzione alla logica formale*. Il Mulino, Bologna.
- [34] Makinson, D.C. (1979). *Temi fondamentali della logica moderna*. Boringhieri, Torino.
- [35] Manca, V. (2001). *Logica matematica*. Bollati Boringhieri, Torino.
- [36] Mangione, C. (1972-1976). *La logica nel ventesimo secolo, I e II Storia del pensiero filosofico e scientifico*, edito da L.Geymonat, tomo VI pagg. 470-682, tomo VII pagg. 303-433. Garzanti, Milano.

- [37] Maracchia, S. (1990). *La storia di un bugiardo, ovvero sia il “paradosso del mentitore”*. La matematica e la sua didattica, 2:59-61.
- [38] Martelli A. e Montanari U. (1982). *An efficient unification algorithm*. ACM Transactions on programming languages and systems, 4:258-282.
- [39] Martin, R.L. (a cura di) (1970). *The paradox of Liar*. Yale University Press.
- [40] Mendelson, E. (1972). *Introduzione alla logica matematica*. Boringhieri, Torino (4a ed.: *Introduction to mathematical logic*. Van Nostrand, Princeton 1997).
- [41] Meschowski, H. (1973). *Mutamenti nel pensiero matematico*. Boringhieri, Torino.
- [42] Monk, J.D. (1972). *Introduzione alla*

teoria degli insiemi. Boringhieri, Torino.

- [43] Nagel, E. e Newman, J.R. (1961). *La prova di Gödel.* Boringhieri, Torino.
- [44] Nathanson, M.B. (1996). *Additive number theory: the classical bases.* Springer, Berlin-Heidelberg-New York.
- [45] Piacentini Cattaneo, G.M. (1996). *Algebra, un approccio algoritmico.* Zanichelli, Bologna.
- [46] Ramsey, F.P. (1964). *I fondamenti della matematica ed altri scritti di logica.* Feltrinelli, Milano (ed. orig.: Proceedings of the London Mathematical Society, 2nd series, 25:338-384, 1925).
- [47] Ribenboim, P. (1989). *The Book of Prime Number Records.* Springer, New York (3a ed.: Springer, New York

1995).

- [48] Rivetti Barbò, F. (1964). *L'antinomia del mentitore nel pensiero contemporaneo da Peirce a Tarski*. Vita e Pensiero, Milano.
- [49] Robinson, J.A. (1965). *A Machine-Oriented Logic Based on the Resolution Principle*. Communications of the ACM, 5:23-41.
- [50] Rogers, R. (1978). *Logica matematica e teorie formalizzate*. Feltrinelli, Milano.
- [51] Rubin, H. e Rubin, J.E. (1963). *Equivalents of the Axiom of Choice*. NorthHolland, Amsterdam.
- [52] Tarski, A. (1969). *Verità e dimostrazione*. Le Scienze, 12 (trad. di M. Servi).
- [53] Tenenbaum, G. e Mendès France, M.

(1997). *Les nombres premiers*.
Presses Universitaires de France,
Paris.

- [54] Turing A. (1936). *On computable numbers, with an application to the Entscheidungsproblem*. Proceedings of the London Mathematical Society, 42:230-265.

Indice analitico

Nota: I link di questo indice si riferiscono alla versione stampata. Negli e-reader, per visualizzare il contenuto a cui l'indice si riferisce, potresti aver bisogno di scorrere in avanti di una o più pagine.

A

- albero, [59](#)
- alfabeto, [68](#)
- algebra di Boole, [75](#)
- antinomia, [45](#)
 - del barbiere, [48](#)
 - del mentitore, [46](#)

di Russell (o degli insiemi norma li),

[47](#)

appartenenza (a un insieme), [4](#)

assioma, [104](#)

assiomi di Peano, [52](#)

assunzione, [110](#)

B

base, [61](#)

C

calcolo, [83](#)

chiusura transitiva, [16](#)

classe di equivalenza, [17](#)

clausola, [158](#)

di Horn, [174](#)

goal, [174](#)

risolvente, [160](#)

unitaria, [159](#)

clausole

contrastanti, 160

genitrici, 160

coerenza, 116

congettura

dei primi gemelli, 65

di Goldbach, 65

congruenza modulo n , 61

connettivi, 91

conseguenza logica, 134

controimmagine, 27

coppia ordinata, 7

costante, 128

generica, 143

crivello di Eratostene, 63

D

differenza, 7

dimostrazione, 104

divisibilità, 63
dominio, 27
di definizione, 30
dualità, 21

E

elemento (di un insieme), 3
enunciato, 90
 atomico, 91
 falsificabile, 95
 insoddisfacibile, 95
 soddisfacibile, 95
 valido, 95
equipotenza, 37
equisoddisfacibilità, 162
equivalenza, 17
equivalenza logica, 94

F

falsificabilità, 95

forma normale

- congiuntiva, 158
- disgiuntiva, 159

formula, 129

- chiusa, 131
- in forma clausale, 159
- in forma prenessa, 162
- indecidibile, 147
- soddisfacibile, 133
- valida, 133
- vera, 133

funzione, 27

- biiettiva, 29
- caratteristica, 44
- composta, 30
- di Skolem, 163
- identità, 32
- iniettiva, 29

parziale, 30
rappresentabile, 146
ricorsiva, 146
suriettiva, 29

I

immagine, 27
induzione, 54
completa, 58
definizioni per, 56
dimostrazioni per, 56
strutturale, 60
insieme, 3
complemento, 7
continuo, 45
delle parti, 6
finito, 39
immagine, 27
infinito, 39

numerabile, 41
quoziente, 18
rappresentazione caratteristica, 4
rappresentazione tabulare, 4
universo, 7
vuoto, 5
insoddisfabilità, 95
interpretazione, 94
intersezione, 6
ipotesi del continuo, 45

L

letterale, 102
linguaggio, 68
linguaggio oggetto, 89

M

massimo, 21
massimo comun divisore, 61

massimo comune minorante, [21](#)
metalinguaggio, [89](#)
minimo, [21](#)
minimo comune maggiorante, [21](#)
modello, [133](#)
Modus Ponens, [109](#)

N

numeri
complessi, [5](#)
composti, [63](#)
di Fibonacci, [59](#)
interi, [5](#)
naturali, [5](#)
primi, [63](#)
razionali, [5](#)
reali, [5](#)
transfiniti, [45](#)

O

ordine, 18

- buon ordinamento, 58

- parziale, 20

- totale, 19

P

porte logiche, 77

potenza

- del continuo, 45

- del numerabile, 41

predicati, 127

preordine, 20

principio

- del terzo escluso, 96

- di contrapposizione, 96

- di non contraddizione, 96

problema della fermata, 148

procedimento euristico, 107

prodotto cartesiano, 8

programmazione logica, 174

proiezione, 8

proprietà

associativa, 73

commutativa, 73

di De Morgan, 76

di idempotenza, 74

distributiva, 73

Q

quantificatore, 129

ambito, 131

esistenziale, 129

universale, 129

R

regola

della doppia negazione, [113](#)
dello Pseudo-Scoto, [113](#)
di calcolo, [176](#)
di consequentia mirabilis, [114](#)
di contrapposizione, [112](#)
di deduzione, [110](#)
di generalizzazione, [142](#)
di inferenza, [104](#)
di ricerca, [176](#)
di scambio delle premesse, [112](#)
di transitività, [112](#)

relazione, [13](#)

- antiriflessiva, [15](#)
- antisimmetrica, [15](#)
- di equivalenza, [17](#)
- di ordine, [18](#)
- di ordine parziale, [20](#)
- di ordine stretto, [18](#)
- di ordine totale, [19](#)

di preordine, 20
inversa, 14
riflessiva, 14
simmetrica, 15
transitiva, 15
reticolo, 75
riduzione all'assurdo, 96
risoluzione, 157
 con insieme di supporto, 173
 lineare, 173
 predicativa, 169
 proposizionale, 158
 sull'ingresso, 173

S

scomposizione in fattori primi, 63
semantica, 84
sintassi, 84
sistema deduttivo

Gentzen predicativo, 141
Gentzen proposizionale, 103
Hilbert predicativo, 142
Hilbert proposizionale, 108
Skolemizzazione, 163
SLD-refutazione, 177
soddisfabilità, 95
sostituzione
 per risposta corretta, 175
 uniforme, 108
sottoinsieme, 6
stringa, 68

T

tableau
 chiuso, 97
 predicativo, 135
 proposizionale, 97
tautologia, 95

tavola di verità, 92
teorema, 104
teoria, 145
 sintatticamente incompleta, 147
termine, 128
 chiuso, 131
 liberamente sostituibile, 132

U

unificatore più generale (mgu), 167
unificazione, 162
 algoritmo, 167
unione, 6

V

validità, 95
valori di verità, 87
variabile, 128
 libera, 131

vincolata, 131