Logica Matematica
Corso di Laurea in Informatica A e B
Università di Torino

Gabriele Lolli

13 marzo 2009

Indice

Presentazione

Lo scopo di questo corso è quello di rendere familiari con le forme di ragionamento tipiche degli argomenti matematici; in informatica in particolare interessano soprattutto quelli che mirano a trovare la soluzione di un problema, a dimostrare che è una soluzione, a presentarla come un algoritmo e a dimostrare che il programma per l'algoritmo è corretto.

Un algoritmo è un insieme articolato e connesso di istruzioni per risolvere un problema; gli algoritmi non sono scritti in un linguaggio di programmazione, ma inizialmente nel linguaggio matematico o addirittura in quello naturale, e in questo devono essere formulati e riconosciuti tali, prima che la loro descrizione guidi alla traduzione nei relativi programmi.

La maggior parte degli algoritmi che sostengono le prestazioni dei calcolatori non sono numerici ma riguardano manipolazioni di simboli (ad esempio l'ordinamento di una lista, o la fusione di due liste in una), quindi la prima consapevolezza — e competenza — da acquisire è che il linguaggio matematico non è solo quello dei numeri, ma abbraccia qualsiasi argomento che si possa riferire ad elementi strutturati.

I ragionamenti relativi devono avere ed hanno lo stesso rigore di quelli numerici, e si svolgono con l'ausilio di un simbolismo appropriato, che è quello della logica matematica (= logica formale moderna).

Gli strumenti che vengono proposti nel corso sono i *linguaggi logici*. Questi hanno una sintassi definita con la stessa precisione e rigidità di quella dei linguaggi di programmazione, ma hanno un carattere universale. Per loro mezzo è possibile rappresentare qualsiasi discorso e ragionamento, con una scelta opportuna dell'alfabeto, nonostante la ridotta dimensione della base logica comune.

Inoltre sono formali, nel senso che inizialmente nessun senso è associato ai simboli dell'alfabeto, ma solo regole di composizione. La loro definizione naturalmente è guidata dalla prospettiva di poterli interpretare su discorsi dotati di senso in qualsiasi dominio di conoscenze, in particolare su quei discorsi che sono dimostrazioni.

In vista della precisione richiesta in matematica (inclusa quella rivolta allo studio degli algoritmi), che non ammette rilassamenti né licenze, è bene realizzare che il concetto di *ragionamento* non è vago, ma si può definire a sua volta in maniera matematica. Ogni ragionamento si può quindi rappresentare in forme standardizzate di passaggi, e un obiettivo del corso è quello di imparare a farlo, usando regole logiche che sono state codificate in vari sistemi.

Il ragionamento stesso, una volta organizzato e rappresentato in modo preciso, si presta a essere oggetto di una trattazione algoritmica — o, fino a che punto questo sia possibile è un problema che ha vari gradi di soluzione, che possono essere illustrate sui due modelli matematici del ragionamento proposti nel corso, quello della logica proposizionale e quello della logica predicativa.

L'argomento degli insiemi è presente ma sussidiario. Il linguaggio insiemistico è usato come integrato nel linguaggio matematico informale, e l'algebra degli insiemi è sviluppata quel tanto che basta per vederne le connessioni con le problematiche logiche e fornire esempi di applicazioni.

Nel testo, il segno di curva pericolosa a margine segnala che si deve prestare particolare attenzione.

Gli esercizi sono di due tipi; alcuni ripetitivi, per soddisfare le richieste di studenti che vogliono tanti esempi sui quali applicare le tecniche o verificare i concetti imparati; gli altri servono per approfondimento e non sono meno importanti, anzi, tutti sono parte del programma.

I riferimenti in nota del tipo "Horstmann, p. 186" rimandano al testo del corso di Programmazione: C. S. Horstmann, *Concetti di Informatica a fondamenti di Java 2*, Apogeo, Milano, 2000.

Le parti scritte in corpo minore sono proposte come letture con informazioni integrative.

Il segno $\square$ è usato per indicare la fine di una dimostrazione, al posto del tradizionale QED, *Quod erat demonstrandum*.

Capitolo 1

Linguaggi

1.1 Dal linguaggio naturale alla logica

La prima competenza che bisogna acquisire è quella della *formalizzazione*, ovvero della traduzione di frasi della lingua naturale o del gergo matematico — che è un misto di formule e di parole — in espressioni di un linguaggio semplificato, schematico e dalla sintassi precisa.

Le frasi che si prendono in considerazione formano un sottoinsieme della totalità delle frasi. Non si considerano espressioni di interrogazione, esclamazione o comando, ma solo frasi dichiarative. Ci si riduce, come primo livello di semplificazione, a frasi elementari che esprimono fatti, e a loro combinazioni mediante particelle logiche.

Non si considerano inoltre frasi con indicatori di tempo e luogo (tempi dei verbi, avverbi di tempo, luogo e modo).

La semplificazione è guidata dalla volontà di restringersi ad espressioni matematiche o comunque preparate alla loro traduzione in programmi (nei linguaggi imperativi si usano comandi, ma questi sono rivolti alla macchina, non costituiscono le frasi da elaborare).

Si devono evitare ambiguità e ridondanze, non per sfizio ma con l'obiettivo di capire e far emergere la struttura logica. Una frase come

La vecchia porta la sbarra

è ambigua perché non è chiara la sua struttura sintattica: se "vecchia" sia un aggettivo sostantivato o un aggettivo, se "porta" e "sbarra" siano sostantivi (nomi) o forme verbali.

Una frase come

Giovanni vede Mario che è malato e piange

e ambigua per ragioni di scansione, occorrono delimitatori come le virgole.

1.1.1 Predicati e relazioni

Le frasi elementari nel linguaggio naturale sono di diverso tipo, ma in tutte si può individuare un soggetto, un verbo e un complemento (eventualmente più soggetti e più complementi, o nessuno). I verbi possono essere intransitivi o transitivi, ed esprimere stati o azioni.

Nella terminologia logica si introducono "predicati" (o "proprietà") e "relazioni"; i primi corrispondono ai verbi intransitivi e alla copula "essere", le seconde ai verbi transitivi.

Si dice che una proprietà è goduta da un soggetto, o che un soggetto ha una determinata proprietà o che soddisfa un predicato. Si dice anche che una proprietà è predicata di un soggetto, espressione dalla quale si vede il collegamento tra i due termini.

Con "la rosa è profumata" o "la rosa profuma" si esprime il fatto che la rosa ha una proprietà, quella di essere profumata. Lo stesso se si dice "la rosa ha profumo". Il verbo "avere" in generale indica possesso, ma non in questo caso. In "Giovanni ama Maria" invece[1] il verbo "amare" ha un soggetto e un complemento oggetto; in logica si dice che sussiste una relazione tra Giovanni e Maria, o che Giovanni e Maria stanno nell'ordine in una relazione, che è la relazione (non simmetrica) di amore.

Tutti i verbi si potrebbero standardizzare nella forma della attribuzione di uno stato a uno o più termini, e questo corrisponderebbe ad avere un solo verbo, la copula "essere", nelle due versioni "essere qualcosa" per i verbi intransitivi e "essere nella relazione ... con" per i verbi transitivi. Questo è il motivo per cui nella trattazione formale successiva (cap. 7) si userà la dizione unica "predicati" per proprietà e relazioni, distinguendo quelli a un argomento (proprietà) da quelli a più argomenti (relazioni). Il "numero di argomenti" è il numero di entità a cui si applica il predicato. Ma informalmente si preferisce distinguere tra predicati in senso stretto (a un solo argomento, o predicati monadici), e relazioni (a più argomenti).

La frase "Giovanni dorme" può diventare "Giovanni ha la proprietà di stare dormendo" (o "Giovanni è addormentato", "Giovanni *sta* dormendo", "Giovanni è nello *stato* di sonno").

"Giovanni possiede un Piaggio 50" diventa "la relazione di possesso sussiste tra Giovanni e un Piaggio 50", o meglio come vedremo "la relazione di possesso sussiste tra Giovanni e una cosa, e questa cosa è un Piaggio 50".

Le frasi matematiche elementari, uguaglianze e disuguaglianze, "è uguale a", "è minore di", rientrano in questa tipologia. Così quelle insiemistiche con "appartiene a", cioè "è un elemento di".

[1] O "Maria è amata da Giovanni", la distinzione tra forma attiva e passiva è inessenziale, salvo che dal punto di vista psicologico.

Alcune frasi possono essere rese sia mediante relazioni che mediante predicati; dipende da come si definiscono le relazioni e i predicati. In "Giovanni è amico di Mario" si può considerare la proprietà "essere amico di Mario" e attribuirla a Giovanni, oppure la relazione "essere amico di" e affermare che sussiste tra Giovanni e Mario.

Non si può dire che una sia giusta e l'altra no; dipende dal contesto; se dopo la prima osservazione si vuole aggiungere che Giovanni piange perché Mario è malato, e bisogna quindi citare di nuovo Mario, si deve usare il nome "Mario" e allora è meglio la versione relazionale, perché in quella con il predicato in nome "Mario" scompare, nella versione formalizzata, assorbito dal simbolo per il predicato: "essere amico di Mario" *in quanto* predicato, nell'analisi logica, è una unità linguistica non scomponibile, anche se espressa in italiano da una successione di parole tra le quali compare "Mario".

Le relazioni a due argomenti, come quelle viste negli esempi, si chiamano binarie. Le relazioni non sono solo binarie: "il punto C giace tra A e B" è un esempio di una relazione ternaria, o tra tre termini.

1.1.2 Termini

I soggetti o gli oggetti, più in generale i termini tra cui sussiste una relazione, sono indicati da vari costrutti linguistici. Il più semplice è il nome proprio, come "Giovanni" e "Maria". Gli altri sono le descrizioni e i nomi comuni.

In "Maria ama il padre di Giovanni", "padre di Giovanni" è una descrizione, ben precisa, di una persona. Analogamente "il quadrato di 2" è una descrizione di un numero; entrambe le descrizioni sono ottenute applicando una funzione[2], nel primo caso "padre di" nel secondo "il quadrato di", a descrizioni più semplici, che in questi esempi sono nomi. Si possono dare descrizioni più complesse, come "la madre del padre di Giovanni" o "meno il quadrato di 2".

I nomi comuni richiedono una trattazione indiretta. Nella frase "Giovanni possiede un Piaggio 50", il Piaggio 50 di Giovanni è uno di una categoria di cose simili; "Piaggio 50" non è un nome proprio, ma un nome comune; è in effetti un predicato, ragione della versione sopra proposta per la frase, che Giovanni possiede una cosa che ha la proprietà di essere un Piaggio 50. Questa frase non è più tuttavia elementare, è in effetti la congiunzione di due frasi: "Giovanni possiede una cosa" e "questa cosa è un Piaggio 50".

Da questo esempio si vede la necessità di chiarire ancora almeno tre aspetti: come rendere "cosa", come rendere la congiunzione delle due frasi, e come

[2]Preciseremo in seguito cosa sono le funzioni dal punto di vista matematico.

rendere “questa cosa”, che nella seconda frase stabilisce un collegamento con la prima.

1.1.3 Connettivi

Le particelle logiche della lingua italiana sono parole come “e”, “oppure”, “se” e altre, che collegano frasi di senso compiuto. Nella lingua italiana queste parole da una parte sono polivalenti e ambigue, hanno diversi sensi — in generale discriminati dal contesto — e dall’altra si presentano in tante versioni equivalenti.

La congiunzione “e” può ad esempio essere resa da una virgola, da “e anche”, da “ma” e ancora altre espressioni. Il senso avversativo di “ma” è uno degli aspetti che vengono lasciati cadere nel passaggio ad un linguaggio formalizzato, in quanto esprime un’aspettativa soggettiva. La congiunzione è resa anche da costrutti più complicati, come “sia ... sia”: “parto sia che piova sia che faccia bel tempo” significa “se fa bel tempo parto, e se piove parto”, magari con l’aggiunta di un “ugualmente” che di nuovo esprime una determinazione soggettiva.

La stessa congiunzione talvolta esprime qualcosa di più o di diverso dalla semplice affermazione di entrambe le proposizioni congiunte; talvolta può significare “e poi”, come in “si sposarono e vissero felici”; talvolta significa “e quindi”, come in “si immerge una cartina di tornasole, e diventa rossa” (se questa frase è intesa non come una descrizione di avvenimenti, nel qual caso “e” significa “e dopo”, ma come come una caratterizzazione di particolari sostanze).

La disgiunzione, “o” o “oppure”, talvolta ha un senso debole (“uno o l’altro o tutt’e due”), talvolta un senso esclusivo (“uno o l’altro ma non tutt’e due”). L’affermazione “piove o c’è il sole” è compatibile con la situazione in cui piove da una nuvola anche se c’è il sole. Il latino aveva due parole diverse *vel* e *aut*, ma la distinzione non è rimasta nelle lingue moderne. Sarà invece ripristinata nei linguaggi formali. La differenza qualche volta è espressa dalla semplice ripetizione di “o” (“o piove o c’è il sole”) ma più spesso dall’enfasi della pronuncia; il tono e il contesto devono essere tenuti presenti per capire il significato inteso. C’è voluto del tempo per tornare a riconoscere due particelle diverse, e anche per accettare *vel* come disgiunzione:

> Alcuni dicono che per la verità di una disgiunzione si richiede sempre che uno dei disgiunti sia falso, perché se entrambi fossero veri non sarebbe una vera disgiunzione, come dice Boezio. Questo però non mi piace. In realtà io dico che se entrambe le parti di

una disgiunzione sono vere, l'intera disgiunzione è vera (Walter Burleigh, *De Puritate*, XCI, 3-19, XIV sec .)

La disgiunzione in italiano talvolta è resa con "ovvero", ma questa parola significa anche "cioè", "vale a dire", cioè una precisazione, non un'alternativa.

La "o" si esprime anche con "altrimenti" come in "Lasciate un messaggio, altrimenti non sarete richiamati", solo apparentemente più ingiuntiva della versione con la "o" (si vedano gli esercizi).

Qualche volta la stessa frase può essere espressa sia con la "e" che con la "o". Si può dire equivalentemente sia "Tutti, bianchi o neri, hanno un'anima", sia "Tutti, bianchi e neri, hanno un'anima". L'affermazione "mele e pere sono frutti" vuole anche dire che "una cosa che sia una mela o una pera è un frutto".

La negazione di una frase si realizza in diversi modi, di solito con la particella "non", inserita però (o soppressa) in vari modi nella frase da negare, con diversi costrutti che coinvolgono altre parole, in particolare i verbi. Da "piove" a "non piove", o "non è vero che piove"; da "qualche volta piove" a "non piove mai"; da "piove sempre" a "qualche volta non piove"; da "non ama nessuno" a "ama qualcuno", da "è bello" a "è brutto", e così via. Per negare "non piove" non si dice "non non piove" ma "piove" o "non è vero che non piove".

Per mettere in evidenza proprietà delle particelle logiche, che non dipendono dal significato delle frasi che connettono, negli esempi proposti useremo d'ora in avanti le lettere A, B, ... per indicare frasi imprecisate, e scriveremo: "A e B", "A oppure B" e simili.

La parola "se" è un'altra particella dai molteplici sensi, e dalle molteplici rese, ad esempio con "B, se A", "A solo se B", "se A allora B", "A implica B", "A, quindi B" — ma "quindi" ha anche un significato temporale, come "poi".

Quando si afferma "se A allora B", A è detta condizione sufficiente per B, e B condizione necessaria per A. "A è condizione sufficiente per B" e "B condizione necessaria per A" sono altri modi di esprimere "se A allora B".

Al "se ... allora" sarà dedicata una discussione speciale per la sua importanza rispetto all'inferenza logica.

Spesso "se ... allora" non è presente in frasi che tuttavia esprimono quel tipo di collegamento: "un numero primo maggiore di 2 è dispari" significa "se un numero è primo e maggiore di 2 allora è dispari". Torneremo sull'argomento.

In considerazione delle ambiguità e molteplicità di espressione messe in luce, un primo passo è quello di introdurre una sola versione fissa delle particelle logiche, sia come simboli che come significati; fatto questo tuttavia, la competenza più importante consiste poi nel saper tradurre le frasi della lingua naturale, disambiguandole quando necessario e possibile, e trovando la versione formale corrispondente.

La precedente discussione non esaurisce certo la complessità della lingua, ma è stata proposta a titolo esemplificativo. Solo una costante (auto)analisi delle varie forme espressive (leggi: tanti esercizi) aiuta a riconoscere le varie insidie.

La standardizzazione è necessaria per poter comunicare con le macchine; ma prima di parlare alle macchine occorre parlare ad altre persone e a se stessi per costruire gli algoritmi. Nell'apprendere a formalizzare si deve anche raffinare la propria logica naturale.

Tuttavia non esiste un elenco completo di quelle che nei linguaggi naturali si riconoscono come particelle logiche. Non abbiamo menzionato ad esempio "né . . . né", o "a meno che".[3] Qualche volta, parole che non sembrano particelle logiche possono essere usate in questo modo, e lo si riconosce nella formalizzazione: "quando" è di solito una determinazione temporale, ma "quando piove, prendo l'ombrello" viene resa quasi necessariamente da "se piove, prendo l'ombrello".

Nell'ottica della formalizzazione, chiedere cosa significa "quando piove, prendo l'ombrello" non è altro che la richiesta di tradurre la frase in un'altra in cui compaia una delle particelle logiche riconosciute tali e scompaia "quando", se non è tra quelle; così si vede a quale delle particelle note la parola è equivalente; ma non sempre è evidente una possibile riduzione di una frase ad un'altra, né sempre una sola.

Esistono peraltro parole anche di difficile catalogazione, che sembrano particelle logiche in quanto legano due frasi, ma hanno sfumature importanti che si perdono nella formalizzazione: ad esempio "siccome piove, prendo l'ombrello", o "prendo l'ombrello perché piove" potrebbero essere espresse dall'asserzione unica "la pioggia è la causa del mio prendere l'ombrello", che coinvolge peraltro la delicata parola "causa"; le frasi contengono tuttavia una determinazione temporale implicita ("sta piovendo"), o anche una qualitativa (un riferimento forse a un particolare tipo di pioggia — a dirotto) che non le rende del tutto equivalenti a "quando piove, prendo l'ombrello" o a "la pioggia è la causa del mio prendere l'ombrello".

Esistono parimenti frasi che ne assommano diverse; la stessa "siccome piove, prendo l'ombrello" invece che una frase può essere considerata un

[3] Si noti l'uso della "o" nella nostra frase, di nuovo scambiabile con "e": si voleva dire che non abbiamo menzionato "né . . . né" *e* non abbiamo menzionato "a meno che"; avremmo potuto dire che non abbiamo *menzionato* né "né . . . né" né "a meno che", *usando* proprio "né . . . né"; l'uso di "o" suggerisce un'altra versione equivalente: "una particella che sia "né . . . né" o "a meno che" non è stata menzionata".

argomento, poiché in essa si afferma un fatto, che piove, oltre a un legame condizionale. Potrebbe corrispondere ad un esempio di *modus ponens* (si vedrà a suo tempo): "Se piove, prendo l'ombrello. Piove. Quindi prendo l'ombrello".

Useremo simboli speciali per rappresentare alcune particelle logiche che sembrano di uso più comune, almeno nei discorsi meno sofisticati. Per queste si potrebbero usare parole della lingua italiana — o comunque di una lingua naturale — fissando per convenzione in modo rigido il loro significato, come si fa ad esempio quando per la congiunzione si usa AND, in informatica. Quando si usano AND e simili, si vuole che il linguaggio sia *friendly* perché ci si deve concentrare su altro; noi invece vogliamo concentrarci proprio su quelle parole, per cui sono meglio simboli nuovi, insoliti, che sorprendano.

Useremo per le particelle logiche i simboli:

$\neg$	per la negazione
$\wedge$	per la congiunzione
$\vee$	per la disgiunzione inclusiva
$\oplus$	per la disgiunzione esclusiva
$\rightarrow$	per il condizionale "se ... allora"
$\leftrightarrow$	per il bicondizionale "se e solo se"

senza escluderne a priori altri, e li chiameremo *connettivi* proposizionali. La negazione è un connettivo unario (cioè agisce su una proposizione), gli altri indicati sono connettivi binari (cioè connettono due proposizioni).

Fissare i simboli è come decidere che in italiano la congiunzione si esprime sempre con"e" e non in altri modi. Per evitare invece la molteplicità di senso occorrerà in seguito dare regole opportune.

La scelta di simboli artificiali è più vantaggiosa anche perché, procedendo, questi simboli non saranno soltanto abbreviazioni, ma insieme ad altri diventeranno una struttura che è essa stessa, se si vuole, oggetto di una teoria matematica, con suoi problemi specifici.

Ad esempio una prima questione, comprensibile anche solo sulla base di quanto detto finora, è se le particelle sopra scelte sono anche fondamentali, e in che senso, o se sono sufficienti, o quante ce ne potrebbero essere. Un'altra riguarda l'equivalenza, affermata per alcuni esempi precedenti, tra frasi diverse espresse con particelle diverse.

Queste strutture forniranno inoltre un ricco campo di scrittura di algoritmi non numerici ma simbolici, applicati a liste o alberi o altre strutture di dati.

Il significato delle particelle logiche è lo stesso a prescindere dal lessico, e per studiarlo occorre non fissarsi su un linguaggio particolare; la trattazione deve valere per tutti, quindi useremo lo stesso artificio matematico di usare lettere per indicare entità non precisate, che nelle applicazioni dovranno essere asserzioni sensate.

La formalizzazione del linguaggio naturale non è qualcosa di meccanico e di compiuto per l'intera gamma delle potenzialità espressive. Esistono argomenti controversi e ancora oggetto di discussioni e di proposte per una formalizzazione soddisfacente - che rientrano in studi più avanzati.

La restrizione alle frasi dichiarative è uno di questi, dal momento che i comandi ad esempio hanno un ruolo apparentemente importante nella programmazione.

Abbiamo visto qualche difficoltà con "siccome" e il suo significato causale. Allo stesso modo è discutibile se "è necessario che . . . " sia da considerare una particella logica: "è necessario che al giorno segua la notte" (o "al giorno segue necessariamente la notte") non sembra equivalente a "al giorno segue la notte" e neanche a "al giorno segue sempre la notte", che è equivalente alla precedente se "segue", privo di determinazioni temporali, assorbe il "sempre"; anche "necessariamente $2+2=4$" forse dice di più di "$2+2=4$", ma non è del tutto chiaro che cosa.

Ancora, è possibile sostenere che il costrutto "è vero che . . . " è pleonastico, in quanto "è vero che piove" è equivalente a "piove", ma è altrettanto possibile sostenere che non è possibile farne a meno.

Altre locuzioni della lingua naturale non formalizzabili le vedremo in seguito.

Esercizi

Esercizio 1.1.1. Esaminare i seguenti discorsi (e altri tratti a scelta da fonti letterarie o giornalistiche) ed individuare le particelle logiche e le frasi elementari (racchiudendole tra parentesi e se necessario riformulando in modo equivalente i discorsi e le loro frasi).

1. Se non è possibile prevedere tutte le azioni delle persone allora o l'universo non è deterministico o le persone non sono perfettamente razionali. Chi sostiene il determinismo deve dunque sostenere che se le azioni delle persone sono prevedibili allora le persone sono perfettamente razionali.

Svolgimento. Introdurre abbreviazioni per le frasi che si ripetono, in modo da arrivare, nel caso del primo brano, a

> Se non Prev allora o non Det o non Raz. Chi sostiene Det allora deve sostenere che se Prev allora Raz

e ancora, togliendo il "chi sostiene", a

> Se non Prev allora o non Det o non Raz. Se Det allora si ha che se Prev allora Raz.

Le abbreviazioni aprono la strada all'uso delle lettere per indicare proposizioni; quando si saranno viste alcune leggi logiche, si potrà tornare a esprimere un giudizio sulla correttezza o meno dell'argomento, che per ora non interessa.

2. Se non è possibile prevedere tutte le azioni delle persone allora o l'universo non è deterministico o le persone non sono perfettamente razionali. Chi sostiene il determinismo deve dunque sostenere che se le azioni delle persone non sono prevedibili allora le persone non sono perfettamente razionali.

3. Se le persone sono interamente razionali, allora o tutte le azioni di una persona possono essere previste in anticipo o l'universo è essenzialmente deterministico. Non tutte le azioni di una persona possono essere previste in anticipo. Dunque, se l'universo non è essenzialmente deterministico, allora le persone non sono interamente razionali.

4. Il numero di queste e di tutte le altre frasi supera il numero dei neuroni del cervello, per cui, anche ammettendo — che non è — che ogni frase richieda un neurone o una combinazione di neuroni per la memorizzazione, non si può pensare che tutte le frasi della competenza linguistica siano immagazzinate in memoria.

Esercizio 1.1.2. Con il costrutto "se ... allora" e le frasi "dico x" e "x è una verità" esprimere: "dico tutta la verità e solo la verità".

Esercizio 1.1.3. Scrivere con le giuste particelle logiche:

1. non c'è fumo senza arrosto
2. fumo vuol dire fuoco.

Esercizio 1.1.4. Come si può esprimere (in almeno due modi)

Lasciate un messaggio, o non sarete richiamati

usando il condizionale?

Esercizio 1.1.5. Trasformare la frase

> Gli studenti che hanno sostenuto la prima *o* la seconda prova di esonero devono ...

nella congiunzione di due frasi.

Esercizio 1.1.6. Trovare altre particelle logiche della lingua italiana, oltre a quelle menzionate nel testo.

Esercizio 1.1.7. Discutere se "cioè" è una particella logica o no, e a quali altre è eventualmente equivalente, in diversi contesti.

Esercizio 1.1.8. Cosa significa per voi "*necessariamente* $2+2=4$"?

1.1.4 Variabili

Gli esempi dell'Esercizio 1.1.1 mostrano che le semplificazioni sono troppo drastiche se si vogliono usare solo i connettivi tra frasi complete; quelle collegate restano complesse e non analizzate. Torniamo perciò a quanto lasciato in sospeso, a come rappresentare "una cosa" e "questa cosa". Nella grammatica, un ruolo fondamentale è svolto dai pronomi, che si presentano in grande varietà, come "uno", "chiunque", "ogni", "qualche" e simili.

I pronomi servono a formare nuove frasi collegandone alcune che hanno un riferimento in comune; nella frase "se uno ha un amico, è fortunato" si individuano due proposizioni componenti "uno ha un amico" e "è fortunato". La seconda frase non presenta il soggetto, ma s'intende che è lo stesso della prima; si può ripetere ("uno è fortunato") oppure più spesso, in altri casi, si deve precisare, con un indicatore che faccia capire esplicitamente che il soggetto è lo stesso (ad esempio "egli", "costui" e simili).

Nella seconda di due frasi collegate, il soggetto della prima può essere presente come oggetto, ad esempio in "se uno è generoso, tutti ne dicono bene", dove "ne" significa "di lui".

Anche per questo tipo di parti del discorso, si hanno molte versioni equivalenti, ciascuna con i suoi vantaggi e la sua convenienza, ad esempio "chiunque abbia un amico è fortunato", "coloro che hanno un amico sono fortunati"; talvolta addirittura basta un'unica frase indecomponibile, come "i generosi sono lodati" per "coloro che sono generosi sono lodati"[4].

è necessario comunque mettere in rilievo il fatto che entrambe le frasi hanno un riferimento comune; se si formalizza la frase "se uno ha un amico, uno è fortunato" introducendo una lettera A per la prima e una lettera B per la seconda, si ottiene $A \to B$ che non mostra la struttura fine della frase, e non permette quindi di indagare se sia vera o no.

[4]La possibilità di questa espressione è all'origine di una diversa analisi del linguaggio, che ha portato alla prima logica formale della storia, nell'opera di Aristotele, come vedremo trattando i sillogismi.

Il simbolismo deve essere arricchito. L'uso dei pronomi è standardizzato per mezzo di simboli che si chiamano variabili: $x, y, \ldots$. Il simbolo x sta per "una cosa", "uno", "una persona" se il discorso si riferisce a esseri umani, "un numero" se il discorso si riferisce ai numeri e così via.

La variabile è creduta un elemento alieno del linguaggio, che compare solo nei simbolismi matematici, ma non è così.

"Se uno ha un amico, è fortunato" equivale nella semiformalizzazione "se x ha un amico, x è fortunato".

Avendo introdotto questi simboli speciali, come peraltro abbiamo già fatto con i connettivi, tanto vale utilizzare anche altre schematizzazioni e completare il distacco del lessico naturale.

Introduciamo perciò simboli per designare predicati, e altri per costruire termini, che corrispondono alle descrizioni.

Useremo preferibilmente

le lettere P, Q, R, ...	per predicati e relazioni
le lettere f, g, ...	per funzioni
le lettere a, b, c, ...	per costanti (nomi propri)
le lettere x, y, ...	con o senza indici, per variabili.

La struttura di una frase del tipo "Giovanni dorme" è rappresentata da "dorme(Giovanni)", o

$$P(a).$$

"Giovanni ama Maria" da "ama(Giovanni, Maria)", o

$$R(a, b).$$

Questa notazione[5] è volutamente analoga a quella delle funzioni, in quanto si pensa che un predicato o una relazione si *applichino* ai soggetti interessati; si potrebbero anche pensare come funzioni, aventi come valori "vero" e "falso".

Più in generale, i termini a cui si applica la relazione non sono necessariamente costanti, o nomi, ma anche descrizioni, come "Il padre di Giovanni ama Maria", che diventa

$$R(f(a), b),$$

o descrizioni incomplete, cioè contenenti variabili, come

$$\text{“Uno dorme”:} \qquad P(x).$$

Tuttavia la rappresentazione grafica scelta per i simboli non è essenziale, per comodità di traduzione si possono anche usare altre lettere, come le

[5]La preciseremo in seguito.

iniziali delle parole italiane (A per “essere amici”), o addirittura complessi di lettere o parole intere, magari in caratteri particolari, come $\textsc{amici}(x, y)$.

Anche la particolare forma $R(a, b)$ non è rigida, talvolta può essere sostituita da $a\ R\ b$. Questo succede in particolare con i simboli per tradizionali relazioni matematiche che hanno adottato tale notazione: $x < y$, $x = y$.

Volgiamoci ora alla formalizzazione della frase “Giovanni possiede un Piaggio 50”, già trasformata sopra in “Giovanni possiede una cosa, e questa cosa è un Piaggio 50”: con una costante g per “Giovanni”, un simbolo di relazione R per “possedere”, un simbolo di predicato P per “Piaggio 50”, si può provare a scrivere

$$R(g, x) \wedge P(x),$$

ma non è sufficiente.

1.1.5 Quantificatori

L’uso delle variabili o della loro versione con pronomi presenta aspetti delicati per trattare i quali il formalismo finora introdotto non è abbastanza discriminante.

Se si dice “A Giovanni piace il Piaggio 50” si intende che a Giovanni piacciono tutti i Piaggio 50, anche se probabilmente desidera averne solo uno (comunque non tutti); se si usa $R(y, x)$ per la relazione “a y piace x” la frase diventerebbe uguale alla precedente, pur avendo un altro senso (in particolare può essere vere o false indipendentemente l’una dall’altra).

Nella frase “se uno ha un amico, è fortunato” ci sono due tipi di “uno”, il primo “uno” è il soggetto, presente tacitamente anche come soggetto di “è fortunato”, e il secondo è l’“un” di “ha un amico”[6]. Il primo “uno” significa “chi”, nel senso di “chiunque”, il secondo significa “qualche”. La stessa parola “uno”, e le corrispondenti variabili x e y possono cioè avere sia un senso universale che uno particolare.

Anche se il senso della frase è ovvio, per chiarezza è meglio dire “*chiunque* abbia *qualche* amico è fortunato”. Così si potrebbe dire “A Giovanni piace un qualunque Piaggio 50” o “A Giovanni piacciono tutti i Piaggio 50” o “A Giovanni piacciono *i* Piaggio 50”. La varietà di costrutti linguistici disponibili nelle lingue naturali ha la funzione di evitare possibili ambiguità in altre frasi di non immediata decifrazione.

Un esempio di frase ambigua, se presa isolatamente, è “uno che segue il corso di Logica si addormenta”. Il professore spera che voglia solo dire che

[6]Non c’è differenza tra “uno” e “un”; si potrebbe dire in entrambi i casi “una persona”, ristabilendo l’uniformità.

si conosce uno studente che tende ad addormentarsi, ma magari gli studenti intendono che tutti si addormentano sempre.

L'uso delle variabili da sole non risolve le ambiguità, anzi le potrebbe accrescere, se vengono a mancare le differenze di significato dei pronomi specifici; in "se x ha y come amico, x è fortunato", se y fosse presa in senso universale, come la x, allora la frase significherebbe che chi è amico di tutti è fortunato, il che è discutibile, piuttosto è un santo.

Un altro esempio è il seguente: nelle due frasi di argomento aritmetico

un numero moltiplicato per se stesso dà 1

e

un numero sommato al suo opposto dà 0

"un numero" è da intendersi in modo diverso; nel primo caso l'unico numero con quella proprietà è 1, e la frase potrebbe essere una sua descrizione estrapolata dal contesto, o un indovinello: "quale è ...?"; nel secondo caso "un numero" significa "qualunque numero".

La differenza non si coglie neanche se si formalizza, la prima frase con $x \cdot x = 1$ e la seconda con $x + (-x) = 0$; per capire la differenza si deve pensare a quali specifici numeri soddisfano le formule, -1 e 1 in un caso, tutti i numeri nell'altro. Nella terminologia usuale, la prima è un'equazione, la seconda un'identità.

Le variabili da sole non rendono la duttilità delle parole che indicano se si parla di uno, qualcuno o tutti.

Si introducono allora due simboli che si chiamano *quantificatori*,

$\forall$ quantificatore universale

e

$\exists$ quantificatore esistenziale

e questi segni si premettono alle formule con variabili per segnalare che, nel loro raggio d'azione determinato dalle parentesi, le variabili stesse devono essere intese nel senso di "tutti" ovvero nel senso di "qualcuno".

La frase "uno che ha un amico è fortunato" diventa, schematizzata,

$$\forall x(\exists y(A(x, y)) \to F(x)).$$

L'uso delle parentesi sarà codificato quando si studierà il formalismo, addirittura in modo pignolo. Per ora basti osservare che quando un quantificatore si riferisce a una variabile fissando il senso delle sue occorrenze, universale o particolare, in una frase, tutta la frase che contiene quelle occorrenze della variabile va racchiusa tra parentesi (nell'esempio, tutta la frase $\exists y(A(x, y)) \to F(x)$ per quel che riguarda x, e $A(x, y)$ per quel che riguarda la y — cosa che è già stata fatta[7]).

Quando si leggono frasi già formalizzate, i quantificatori $\forall x$ e $\exists x$ si leggono usualmente sempre nello stesso modo: "per tutti gli x" (o "per ogni x") e "esiste un x tale che" (o "esistono x tali che"), anche quando non è la lettura più elegante. Invece in italiano ci sono diversi modi di esprimersi.

Alcune espressioni della lingua naturale hanno tuttavia significati colloquiali che non hanno interesse logico e che comunque non sono esprimibili nel formalismo. Anzi bisogna fare attenzione a non lasciarsi influenzare. Ad esempio "qualche" viene spesso usato per dire "pochi", per indicare un certo numero ma non grande, e spesso maggiore di uno, ché se è uno si dice "uno". Invece $\exists x$ vuol sempre dire "esiste almeno un ...", e possono essere uno, dieci o centomila, o anche tutti.

Quando si usa "qualche", talvolta in italiano si sottintende "non tutti"[8]; invece $\exists x \ldots$ è compatibile col fatto che tutti soddisfino la condizione; è solo un'affermazione più debole: se si sa che tutti i gamberi sono rossi, si può affermare $\exists x(\text{gambero}(x) \wedge \text{rosso}(x))$ come vero; naturalmente così non si afferma che tutti i gamberi sono rossi (che sarebbe reso da $\forall x(\text{gambero}(x) \to \text{rosso}(x)))$ ma che esiste un gambero rosso.

Le variabili svolgono il ruolo di "uno", "una cosa", "un numero" e simili; di quale esattamente dipende dall'universo di discorso. Questo va precisato, in vari modi. Spesso la scelta dei predicati e delle relazioni suggerisce implicitamente di cosa si parla: se si usa una relazione A per "essere amico di ..." è implicito che si parla di persone o animali. Allora $\forall x(\exists y A(x, y) \to F(x))$ si legge "ogni persona o animale che abbia ...".

Tuttavia è difficile che il discorso entro il quale si inserisce $\forall x(\exists y A(x, y) \to F(x))$ si limiti a persone o animali; nel prosieguo possono essere menzionate anche cose o idee. Al di fuori della matematica, dove è di solito ben precisato,[9] l'universo di discorso è ricco e variegato. $\forall x \ldots$ si legge dunque "per ogni x ..." dove x a priori può stare per gli elementi più disparati.

[7] In realtà dopo $A(x, y)$ la y non occorre più e non c'è bisogno delle parentesi per una lettera corretta, come sarà spiegato in seguito: scriveremo anche $\forall x(\exists y A(x, y) \to F(x))$.

[8] Da un compito in classe: "Se qualche triangolo isocele è equilatero, di conseguenza qualche triangolo isocele non lo è". La conclusione è vera, ma "di conseguenza" no, e l'unico modo per immaginare come sia stata concepita è l'interpretazione di "qualche" come "non tutti".

[9] Non sempre: s e si discute una equazione e non si precisa quale è il dominio numerico, le risposte possono essere bene diverse.

In molte frasi tuttavia i quantificatori chiaramente non si riferiscono a tutti gli elementi dell'universo di discorso ma a parti più ristrette; le frasi aritmetiche per esempio raramente iniziano con "tutti i numeri", piuttosto con "tutti i numeri positivi", o "tutti i numeri primi"; e raramente si parla di tutti gli esseri viventi, ma piuttosto di tutti gli uomini, o di tutte le donne, o di tutti gli italiani e così via restringendo.

Nel formalismo logico la restrizione dei quantificatori avviene nel seguente modo. La frase "tutti i tedeschi sono biondi" si rappresenta con due predicati, "tedesco" e "biondo", e la forma

$$\forall x(T(x) \to B(x)),$$

dove il quantificatore $\forall x$ è letto "per tutte le persone", cioè con la x che varia su tutto l'universo del discorso (la specie umana): "per ogni x, se x è tedesco allora x è biondo.

Questa forma è corretta grazie alle proprietà del condizionale, che vedremo meglio in seguito. Se $T(x) \to B(x)$ è vero per tutte le persone, allora ogni tedesco rende vero il condizionale, l'antecedente e quindi vero il conseguente, ed è vero che tutti i tedeschi sono biondi; se viceversa è vero che tutti i tedeschi sono biondi, anche l'enunciato di sopra che si riferisce con $\forall x$ non ai tedeschi ma a tutte le persone è vero: se uno è tedesco, allora è biondo e il condizionale è vero; se Giovanni è brutto ma non è tedesco, lo si vorrà considerare un controesempio che falsifica l'affermazione? Non sembra ragionevole; si assume che $T(\text{Giovanni}) \to B(\text{Giovanni})$ sia vero, e così $T(x) \to B(x)$ è vera per tutte le persone.

In pratica, gli aggettivi sono resi da predicati con l'ausilio del condizionale: in "tutte le persone tedesche sono bionde" l'aggettivo "tedesco" diventa il predicato "essere tedesco" e la frase "tutte le persone, se sono tedesche, sono bionde".

"Tutti i P sono $\ldots$" e "qualche P è $\ldots$", dove P delimita il campo di variabilità del riferimento, si realizzano dunque introducendo un predicato unario P e scrivendo rispettivamente $\forall x(P(x) \to \ldots)$ e $\exists x(P(x) \wedge \ldots)$. Si noti ovviamente la differenza nel caso del quantificatore esistenziale, dove la restrizione è realizzata con la congiunzione, che viene dalla traduzione di "esiste uno che è P e che $\ldots$".

In particolare è da sottolineare che si usa un solo tipo di variabili; nella pratica matematica talvolta se ne usa più di uno, ad esempio in geometria lettere maiuscole $A, B, \ldots$ per punti e minuscole $r, s, \ldots$ per rette. Ma ci si riconduce a un solo tipo di variabili usando gli opportuni predicati, ad e sempio "essere un punto" e "essere una retta".

1.1.6 Esempi:

dal linguaggio naturale...

Esempio 1.1.9. "Maria ama il padre di Giovanni" è formalizzata da

$$A(m, f(g)),$$

dove m e g sono costanti, m per "Maria" e g per "Giovanni", ed f un simbolo funzionale per "il padre di ...".

Esempio 1.1.10. Per formalizzare "Maria ama il figlio di Giovanni" non si può usare un simbolo f per "il figlio di", perchè "figlio di" non è una funzione univoca: a una persona possono corrispondere diversi figli, o nessuno. Allora "Maria ama il figlio di Giovanni" si formalizza come sotto "Maria ama un figlio di Giovanni" e a parte si afferma che Giovanni ha un solo figlio (vedremo come).

Esempio 1.1.11. "Maria ama un figlio di Giovanni" è formalizzata da

$$\exists x(A(m, x) \land F(x, g)),$$

letta

esiste un x tale che Maria ama x e x è figlio di Giovanni,

dove F è un simbolo *relazionale* a due posti, e $F(x, y)$ sta per "x è figlio di y".

Si potrebbe anche dire "Maria ama uno, che è figlio di Giovanni" o "Maria ama un tizio che è figlio di Giovanni".

In questo caso "figlio di Giovanni" ha la funzione di nome comune, come "Piaggio 50" in "Giovanni possiede un Piaggio 50", e infatti si formalizza nello stesso modo.

Esempio 1.1.12. "Maria ama i figli di Giovanni", che significa che Maria ama tutti i figli di Giovanni, si formalizza con

$$\forall x(F(x, g) \to A(m, x))$$

e non con $\forall x(A(m, x) \land F(x, g))$; questa significa che tutti sono figli di Giovanni, e che Maria li ama tutti; il che implica che Giovanni sia Dio, e forse Maria la Madonna.

Per la formalizzazione corretta, può essere utile vedere nella frase un caso di quantificatore ristretto, ai figli di Giovanni, leggendola "Tutti i figli di Giovanni, Maria li ama" o al passivo: "Tutti i figli di Giovanni sono amati da Maria".

Esempio 1.1.13. "Non tutte le ciambelle riescono col buco".

Si scelga un predicato C per "essere una ciambella" e una relazione $B(x, y)$ per "x è un buco di y". Quindi si trasforma la frase eliminando "non tutte" a favore di "qualche ciambella non riesce col buco". "Riuscire con buco" o "avere il buco" possono essere trattate come equivalenti: la prima versione allude al processo di fabbricazione che finisce male, la seconda al risultato. Allora

$$\exists y(C(y) \land \neg\exists x B(x, y)).$$

Esempio 1.1.14. "Ogni rosa ha le sue spine".

Sia R il predicato "essere una rosa" e $S(x, y)$ la relazione "x è una spina di y".

$$\forall x(R(x) \to \exists y S(y, x)).$$

Si noti che se $S(x, y)$ è la relazione "x è una spina di y", $S(y, x)$ si legge "y è una spina di x". C'è grande libertà nell'uso delle variabili: $\exists y S(y, x)$ si potrebbe scrivere anche $\exists z S(z, x)$; in italiano hanno la stessa traduzione "x ha qualche spina". Quello che importa è non usare la stessa variabile quando devono essere distinte: se si scrive $\exists x S(x, x)$ si dice che c'è una rosa che è una spina.

Esempio 1.1.15. "Ogni rosa ha qualche spina".

La frase è la stessa di prima, perché se una rosa ha delle spine queste sono sue. Entrambe possono comunque essere formalizzate anche in un altro modo, con un predicato per "essere una spina" e una relazione binaria H di possesso:

$$\forall x(R(x) \to \exists y(S(y) \land H(x, y))).$$

Esempio 1.1.16. "Chi rompe paga e i cocci sono suoi".

"Rompere" è verbo transitivo, salvo che in usi metaforici, quindi bisogna pensare che si dica "chi rompe qualcosa, una qualunque cosa"; sia $R(x, y)$ la relazione "x rompe y"; sia quindi $C(x, y)$ la relazione che intercorre tra due pezzi di materia se il primo è un coccio dell'altro, cioè un pezzo che risulta dalla sua rottura; scegliamo la relazione $S(x, y)$ a indicare che x paga il valore di y, e sia infine $H(x, y)$ la relazione "x assume il possesso di y". Allora

$$\forall x \forall y(R(x, y) \to S(x, y) \land \forall z(C(z, y) \to H(x, z))).$$

Il complesso $\forall x \forall y \ldots$ si legge "per ogni x e per ogni y ...". È anche lecito abbreviare con $\forall x, y \ldots$, così come $\exists x \exists y \ldots$ con $\exists x, y \ldots$.

"Chiunque rompa qualunque cosa ..." o "Qualunque cosa uno rompa ..." sono equivalenti: in base a questa lettura è evidente che risulterà che $\forall x \forall y \ldots$ è equivalente a $\forall y \forall x \ldots$ e che $\exists x \exists y \ldots$ è equivalente a $\exists y \exists x \ldots$.

La precedente formula è tuttavia ambigua, e deve essere corretta in

$$\forall x \forall y (R(x, y) \to (S(x, y) \land \forall z (C(z, y) \to H(x, z))))$$

in modo che entrambe le conseguenze (pagare e tenere i cocci) dipendano da $R(x, y)$. Altrimenti se si pensasse a $(R(x, y) \to S(x, y)) \land \ldots$ la frase $\forall z (C(z, y) \to H(x, z))$ significherebbe che x si prende i cocci di ogni cosa, che l'abbia rotta lui o no.

Esempio 1.1.17. "Un regalo conquista un amico".

Cominciamo a riformulare la frase spogliandola di significati metaforici (un regalo è una cosa e non conquista nulla). Si intende ovviamente dire che chi fa un regalo acquista un amico, e più dettagliatamente che se una persona fa un regalo a un'altra persona, questa diventa suo amico. Usiamo una relazione ternaria $R(x, y, z)$ per "x regala y a z" e una relazione binaria per $A(x, y)$ "x diventa amico di y".

$$\forall x \forall y (\exists z R(x, z, y) \to A(y, x)).$$

Esempio 1.1.18. "A Natale si fanno regali agli amici".

Si intende che a Natale ognuno fa un regalo a ciascuno dei suoi amici. Non è il caso di mettere in evidenza "Natale", che non è rilevante per la struttura logica della frase. Usiamo una relazione ternaria $R(x, y, z)$ per "x a Natale regala y a z" e una relazione binaria $A(x, y)$ per "y è un amico di x".

$$\forall x \forall y (A(x, y) \to \exists z R(x, z, y)).$$

Esempio 1.1.19. "Chi non risica non rosica".

"Risicare" è verbo intransitivo (anche se qualche volta si dice "ha rischiato qualcosa", ma si intende "ha rischiato un po' "). "Rosicare" è transitivo, anche se nella frase non compare il complemento oggetto, ma si intende "non rosica nulla". Usiamo un predicato R per "risicare" e una relazione $S(x, y)$ per "x rosica y".

$$\forall x (\neg R(x) \to \neg \exists y S(x, y)).$$

Esempio 1.1.20. "Sono eligibili tutti e soli gli studenti in corso".

Non interessa a cosa siano eligibili; serve un predicato per "essere eligibile", uno per "essere studente" e uno per "essere in corso".

$$\forall x (E(x) \leftrightarrow S(x) \land C(x)).$$

La dizione "tutti e soli" è strettamente legata a "se e solo se". "Tutti gli studenti in corso sono eligibili" è formalizzata da

$$\forall x (S(x) \land C(x) \to E(x)),$$

mentre “solo gli studenti in corso sono eligibili” da

$$\forall x(E(x) \to S(x) \land C(x)).$$

La congiunzionne di queste due ultime frasi è equivalente, come vedremo, alla prima.

... dalla matematica

Esempio 1.1.21. La frase “dati due numeri, uno minore dell’altro, esiste un terzo numero compreso tra i due”, vera nel campo dei razionali e in quello dei reali, falsa negli interi, può essere resa da

$$\forall x \forall y(x < y \to \exists z(x < z \land z < y)).$$

La congiunzione $x < z \land z < y$ si può abbreviare, secondo l’uso matematico, con $x < z < y$.

Non esiste un quantificatore che quantifichi sulle coppie; ci si comporta come se la frase fosse “dato un primo numero e dato un secondo numero ...”. Ma “un primo” e “un secondo” servono solo a facilitare l’espressione, si sarebbe potuto dire anche “dato un numero e dato un numero ...”, con qualche difficoltà nel seguito per i riferimenti appropriati.

Si faccia attenzione che neanche la presenza di “due” vuol dire che i numeri devono essere considerati diversi; tale forma comune di espressione distingue il modo, il momento in cui i numeri sono presentati, o pensati, ma non è escluso in generale che si presenti lo stesso numero due volte.

Nell’esempio 1.1.18 precedente, a Natale uno fa anche regali a se stesso, se si vuole bene.

“Dati due numeri” significa “fatta due volte la scelta di un numero”, e le scelte possono cadere sullo stesso numero. In termini probabilistici, si tratta di scelte con reimmissione; oppure si deve considerare che la scelta di un numero non lo toglie certo dall’insieme. “Dati due numeri, esiste la loro somma” si può scrivere

$$\forall x \forall y \exists z(z = x + y)$$

ma esiste anche la somma di ogni numero con se stesso; x e y possono prendere tutti i valori in tutte le combinazioni possibili, quindi anche valori uguali.

Quando tuttavia si mette come sopra la condizione “uno minore dell’altro” — come nella frase proposta — allora si esclude che possano essere uguali perchè la relazione “minore di” non è riflessiva. Tuttavia lo si esclude solo attraverso una deduzione, non con la semplice scrittura: se x e y denotano lo stesso numero, e bisogna considerare anche questo caso per verificare se la frase è vera, in $x < y \to \exists z(x < z \land z < y)$ l'antecedente $x < y$ risulta falso (come nell'esempio dei tedeschi).

Con "un terzo" di nuovo si vuol dire semplicemente "un numero", e che sia diverso dai primi due segue automaticamente se "compreso" significa "strettamente compreso"; altrimenti, se la relazione d'ordine fosse intesa come $\leq$ allora potrebbe anche essere uguale a uno dei due; non è questo il senso della frase, che vuole esprimere la densità dell'ordine dei numeri reali — e anche dei razionali.

Se nella stessa formula il segno di relazione è interpretato su di una relazione riflessiva, come

$$\forall x \forall y(x \leq y \to \exists z(x \leq z \land z \leq y)),$$

o più in generale "se R è riflessiva allora ...", ovvero

$$\forall x R(x,x) \to \forall x \forall y(R(x,y) \to \exists z(R(x,z) \land R(z,y))),$$

allora la formula è banalmente vera per ogni relazione[10].

Esempio 1.1.22. "La relazione R è riflessiva", che significa che ogni elemento sta nella relazione R con se stesso, si scrive

$$\forall x R(x,x),$$

come abbiamo fatto sopra.

Esempio 1.1.23. "La relazione R è simmetrica", che significa che se la relazione R sussiste tra uno primo e un secondo elemento allora sussiste anche tra il secondo e il primo, si scrive

$$\forall x \forall y(R(x,y) \to R(y,x)).$$

Esempio 1.1.24. "La relazione R è transitiva", che significa che se R sussiste tra un primo elemento e un secondo, e tra questo e un terzo, allora sussiste anche tra il primo e il terzo, si scrive,

$$\forall x \forall y \forall z(R(x,y) \land R(y,z) \to R(x,z)).$$

Esempio 1.1.25. Come non esiste un quantificatore sulle coppie, così non esiste un quantificatore che esprima "esiste esattamente un ...", o "esiste un solo ...". Tale locuzione si realizza mediante l'uguaglianza come nel seguente esempio.

La frase "dati due numeri, esiste un solo numero che è la loro somma" si formalizza come

$$\forall x \forall y \exists z(z = x + y \ \land \forall u(u = x + y \to u = z)).$$

[10] Con "banalmente" s'intende che dati x e y come z si può prendere o x o y, e la formula non ci dà veramente informazioni.

In generale "Esiste un solo x tale che $P(x)$" si formalizza come

$$\exists x(P(x) \wedge \forall y(P(y) \to x = y)).$$

Esempio 1.1.26. In modo analogo si può esprimere la locuzione "esistono esattamente due elementi tali che ..." (esercizio).

Suggerimento. Si scriva prima "esistono almeno due elementi tali che ...", ricordando quanto detto nell'esempio 1.1.21 a proposito delle coppie di quantificatori.

Esempio 1.1.27. Non si riesce invece con nessun giro di formule del formalismo che stiamo usando ad esprimere "la maggior parte degli elementi ..." o "quasi tutti ...".

Esempio 1.1.28. Analogamente non si riesce ad esprimere "tanti".

La frase "dati due numeri diversi tra loro, esiste un numero che è propriamente compreso tra i due numeri dati" si rappresenta con

$$\forall x \forall y(x \neq y \to \exists z(x < z < y \vee y < z < x)).$$

Esempio 1.1.29. La frase "ogni numero positivo ha una radice quadrata", vera nei reali, falsa nei razionali, si rappresenta come

$$\forall x(0 < x \to \exists y(x = y^2)),$$

dove con y^2 si indica la funzione potenza di esponente 2.

Esempio 1.1.30. "Un numero è divisibile per un altro numero se e solo se esiste un terzo numero che moltiplicato per il secondo dà il primo".

Scriviamo $x|y$ per "y è divisibile per x" o "x divide y" e usiamo il solito segno di moltiplicazione:

$$\forall x \forall y(x|y \leftrightarrow \exists z(y = x \cdot z)),$$

ma di nuovo si noti che x, y, z non devono necessariamente indicare numeri tutti diversi tra loro.

Esempio 1.1.31. "Esistono due numeri primi consecutivi".

Per questa frase complicata procediamo in due passi; usiamo un'abbreviazione $\mathrm{pr}(x)$ per "x è primo " e scriviamo

$$\exists x \exists y(x = y + 1 \wedge \mathrm{pr}(x) \wedge \mathrm{pr}(y))$$

riservandoci di sostituire $\mathrm{pr}(x)$ con la sua scrittura corretta data nel prossimo esercizio.

Che i numeri siano due non risulta dallo scrivere $\exists x \exists y$ ma da $x = y + 1$ che implica $x \neq y$ (lo si deduce facilmente dagli assiomi dei numeri naturali); si potrebbe anche scrivere:

$$\exists x(\mathrm{pr}(x) \wedge \mathrm{pr}(x+1)),$$

dando per scontato, come sopra, che $x \neq x + 1$.

Esempio 1.1.32. "Un numero è primo se e solo se è maggiore di 1 ed è divisibile solo per 1 e per se stesso".

Per esprimere questa che è la definizione di un nuovo predicato usiamo un nuovo simbolo $\mathrm{pr}(x)$ e scriviamo

$$\forall x(\mathrm{pr}(x) \leftrightarrow x > 1 \wedge \forall z(z|x \rightarrow z = 1 \vee z = x))$$

Esempio 1.1.33. "2 è l'unico numero primo pari".

"Numero pari" significa "divisibile per 2". La frase si può trasformare in "2 è primo e pari e se un numero è primo e pari allora è uguale a 2". Quindi

$$\mathrm{pr}(2) \wedge 2|2 \wedge \forall x(\mathrm{pr}(x) \wedge 2|x \rightarrow x = 2).$$

Esempio 1.1.34. "3 è dispari"

Il predicato "dispari" si può definire come "non pari" e quindi

$$\neg 2|3,$$

o meglio $\neg(2|3)$ perché $\neg$ non si confonda con un segno aritmetico[11], oppure dicendo che un numero dispari è della forma $2 \cdot y + 1$, e in aritmetica si dimostra che le due definizioni sono equivalenti, quindi

$$\exists y(3 = 2 \cdot y + 1).$$

Esempio 1.1.35. "Ogni primo maggiore di 2 è dispari"

è un caso di quantificatore ristretto, ma lo si può restringere in due modi: ai numeri primi oppure ai numeri primi maggiori di 2. Il predicato "essere primo maggiore di 2" si può definire con $(\mathrm{pr}(x) \wedge x > 2)$ e si ha allora, se si scrive $\mathrm{disp}(x)$ per "x è dispari ",

$$\forall x((\mathrm{pr}(x) \wedge x > 2) \rightarrow \mathrm{disp}(x)).$$

Oppure se si restringe solo ai primi si deve scrivere

$$\forall x(\mathrm{pr}(x) \rightarrow (x > 2 \rightarrow \mathrm{disp}(x))).$$

In questo caso le parentesi interne servono a evidenziare la composizione corretta della frase mediante le due occorrenze del condizionale.

[11] Che non si legga $\neg 2$ come "diverso da 2".

Esempio 1.1.36. "Esistono numeri pari arbitrariamente grandi".

La locuzione "arbitrariamente grandi" o "grandi quanto si vuole" significa che comunque si dia un numero, ne esiste uno più grande con la proprietà in oggetto — non che un numero è grande quanto si vuole, un numero è quello che è. Quindi

$$\forall x \exists y (x < y \wedge 2 | y).$$

Esempio 1.1.37. "Ci sono almeno due quadrati minori di 10".

Consideriamo 10 una costante (in realtà è un termine complesso). "x è un quadrato" significa che x è il quadrato di qualche numero, e si formalizza come $\exists u(x = u^2)$. Quindi

$$\exists x \exists y (x \neq y \wedge x < 10 \wedge y < 10 \wedge \exists u(x = u^2) \wedge \exists u(y = u^2)),$$

dove $x \neq y$ è un'abbreviazione per $\neg(x = y)$.

Si noti che da $\exists u(x = u^2) \wedge \exists u(y = u^2))$ non segue che la u sia la stessa, e quindi x e y uguali; le due frasi sono indipendenti; è come se si dicesse: "esiste *un numero* il cui quadrato è x ed esiste *un numero* il cui quadrato è y"; non vuol dire che sia lo stesso numero. Ma si sarebbe potuto anche scrivere $\exists u(x = u^2) \wedge \exists v(y = v^2))$.

Esempio 1.1.38. "Per due punti passa una e una sola retta".

Primo modo. Usiamo variabili diverse per punti e rette e una relazione binaria Q per "un punto giace su una retta".

$$\forall A \forall B \exists r (Q(A, r) \wedge Q(B, r) \wedge \forall s (Q(A, s) \wedge Q(B, s) \to r = s))$$

Secondo modo. Usiamo un solo tipo di variabili e due predicati P per "essere un punto" e R per "essere una retta".

$$\begin{aligned} \forall x \forall y (P(x) \wedge P(y) \to \exists z (R(z) \wedge Q(x, z) \wedge Q(y, z) \wedge \\ \forall u (R(u) \wedge Q(x, u) \wedge Q(y, u) \to z = u)). \end{aligned}$$

Le due soluzioni sono equivalenti; nella prima si usa un linguaggio a due sorta di variabili, che ha le stesse proprietà logiche di quello con una sola sorta.

Esempio 1.1.39. "La funzione $y = x^3$ è iniettiva e suriettiva" si formalizza

$$\forall x_1 \forall x_2 (x_1 \neq x_2 \to x_1^3 \neq x_2^3) \wedge \forall y \exists x (y = x^3),$$

in un linguaggio che abbia il simbolo $=$ e un simbolo funzionale indicato con x^3.

Esempio 1.1.40. L'affermazione che la relazione "$y = 2 \cdot x$" è una relazione funzionale e iniettiva è formalizzata da:

$$\forall x \exists y(y = 2 \cdot x \wedge \forall z(z = 2 \cdot x \to z = y)) \wedge \forall x_1 \forall x_2(x_1 \neq x_2 \to 2 \cdot x_1 \neq 2 \cdot x_2).$$

Dagli esempi si traggono diverse regole euristiche: riformulare la frase in un italiano semplice, con soggetto, verbo e complementi; trasformare i pronomi quantitativi come "ognuno", "alcuni", "nessuno", "uno" ... usando sempre solo "per tutti ..." ed "esiste un ...", anche se la frase diventa barocca; guardare i verbi, se sono intransitivi o transitivi, e sostituire le frasi elementari con le dizioni "ha la proprietà ..." e "sussiste la relazione ..."; non prendere relazioni troppo inglobanti che nascondano la sintassi informale, immaginando possibili proseguimenti della frase che richiedono di riprendere certi elementi; invece lasciare cadere particolari empirici; nelle frasi matematiche, risalire sempre alle definizioni dei termini coinvolti.

1.1.7 Esercizi

Esercizio 1.1.41. Formalizzare frasi del linguaggio comune come le seguenti e altre a piacere:

1. Il mio cellulare è migliore del tuo ma costa di più.
2. Chi è senza peccato scagli la prima pietra.
3. Senza soldi si vive male.
4. Senza amici si è soli.
5. Un giorno sì e uno no piove.
6. I supermercati abbondano di ogni ben di Dio.
7. Maria ama il figlio di Giovanni.
8. Il 33, 3% circa degli italiani possiede due macchine.
9. Chi ha superato solo una delle due prove di esonero, all'esame porta solo la parte non superata.
10. Chi lascia la via vecchia per la nuova sa quel che lascia ma non sa quel che trova.

Esercizio 1.1.42. Formalizzare "A Giovanni piacciono i Piaggio 50"; trasformarla prima in italiano in modo da far comparire un quantificatore ristretto.

Esercizio 1.1.43. Formalizzare “I professori premiano gli studenti meritevoli”.

Esercizio 1.1.44. In un'assemblea di politici, questi si dividono in onesti e disonesti, e si sa che

a) esiste almeno un politico onesto;

b) presi due politici a caso, uno almeno è disonesto.

Si formalizzino le condizioni sui politici.

Se nell'assemblea ci sono cento politici, si può decidere (dedurre) quanti sono gli onesti e quanti i disonesti? Formalizzare anche la risposta.

Esercizio 1.1.45. Usando un predicato S per “x è uno studente”, una relazione C per “x è un computer”, un predicato F per “x è funzionante” e una relazione U per “x utilizza y”, formalizzare le seguenti frasi:

(a) Un computer non è utilizzato da nessun studente.

(b) Ogni computer funzionante è utilizzato da almeno uno studente.

(c) Non tutti i computer sono funzionanti.

Quale dei seguenti enunciati è una traduzione corretta di (a)?

$$\exists x(C(x) \wedge \forall y(\neg S(y) \wedge \neg U(y,x)))$$
$$\exists x(C(x) \rightarrow \forall y(S(y) \rightarrow \neg U(y,x)))$$
$$\exists x(C(x) \wedge \forall y(S(y) \rightarrow \neg U(y,x)))$$

Qual'è il significato delle altre formule? Tradurre (b) e (c) in opportuni enunciati del linguaggio formale introdotto.

Esercizio 1.1.46. Una base di dati relazionale è una tabella come la seguente:

Codice genitori		Nome figli
Codice	1	Paola
Codice	1	Francesco
Codice	2	Paola
Codice	3	Anna
Codice	3	Lea
Codice	4	Giovanni

Le interrogazioni sono del tipo: "(insieme dei) codici dei genitori che non hanno figli di nome Paola". Se si usa una relazione $R(x,y)$ per "x è il codice di genitori che hanno figli di nome y", rappresentata dalla tabella, l'interrogazione si formalizza come $\neg\exists y(R(x,y) \land y = \text{“Paola”})$, che definisce (vedi capitolo 2) l'insieme $\{x : \neg\exists y(R(x,y) \land y = \text{“Paola”})\}$.

Formalizzare la interrogazione "codici dei genitori che hanno figli di nome diverso da Paola".

1.2 Deduzione naturale

Le frasi non sono mai pronunciate in modo isolato, ma sono collegate tra loro in discorsi. Questi, quando non siano chiacchiere o racconti, ma ragionamenti, portano da certe frasi a certe conclusioni.

Anche i ragionamenti ammettono una standardizzazione, almeno quelli di occorrenza più frequente. La forma standard dei ragionamenti che consideriamo, e che sono quelli che si incontrano in matematica[12], si chiama deduzione.

Un ragionamento lo si può vedere come una trasformazione di una frase, in un'altra, attraverso una serie di tappe intermedie. Le tappe intermedie si chiamano inferenze.

Abbiamo già incontrato alcuni casi in cui abbiamo detto che due frasi diverse dicono la stessa cosa, o che una frase ammette diverse rappresentazioni. In questi casi per giustificare l'affermazione occorrono due ragionamenti, uno dalla prima a concludere la seconda, l'altro viceversa.

Le inferenze più comuni, la cui composizione e combinazione produce le deduzioni, si appoggiano a proprietà e condizioni di uso delle particelle logiche, ovvero dei connettivi e dei quantificatori.

Ad esempio chiunque accetterebbe come corretto il ragionamento incluso nel seguente dialogo, dovunque porti il seguito:

— Hai letto *Guerra e pace* e *I promessi sposi*?

— Sì.

— Dunque, siccome hai letto *I promessi sposi* saprai ...

Analogamente chiunque assentirebbe a questa inferenza:

— Hai letto *Guerra e pace*?

[12] Non prendiamo in considerazione i ragionamenti basati su induzioni empiriche, o su valutazioni soggettive di probabilità.

— Sì.

— Hai letto *I promessi sposi*?

— Sì.

— Dunque, siccome hai letto *Guerra e pace* e *I promessi sposi* mi sai dire …

In entrambi i casi, la conclusione, individuata da “dunque” è intrinsecamente legata al significato della congiunzione.

I due tipi di inferenza si schematizzano nel seguente modo:

(E$\wedge$) $$\frac{A \wedge B}{A} \qquad \frac{A \wedge B}{B}$$

che è una regola, sdoppiata in due per evidenti ragioni di simmetria, di eliminazione della congiunzione, e

(I$\wedge$) $$\frac{A \quad B}{A \wedge B}$$

che è la regola di introduzione della congiunzione.

Sopra alla riga orizzontale stanno le premesse, o la premessa, della regola, da intendere date in modo non ordinato, sotto la riga la conclusione.

Le regole permettono di codificare le inferenze elementari sopra portate ad esempio.

In modo analogo, facendo riferimento a usi comuni, si riconoscono corrette le seguenti regole di eliminazione della disgiunzione:

(E$\vee$) $$\frac{A \vee B \quad \neg A}{B} \qquad \frac{A \vee B \quad \neg B}{A}$$

e di introduzione della disgiunzione

(I$\vee$) $$\frac{A}{A \vee B} \qquad \frac{B}{A \vee B}$$

entrambe sdoppiate in due per evidenti ragioni di simmetria.

Due esempi molto semplici di uso implicito di queste inferenze sono i seguenti.

1. Il numero 2, come tutti i numeri naturali, è ≥ 0; ma $2 \neq 0$ quindi $2 > 0$.

2. Se piove o nevica non esco. Siccome piove, non esco.

   La conclusione è mediata da una inferenza tacita, omessa, cioè che "se piove, allora piove o nevica".

L'ultimo ragionamento si appoggia a una regola tipica del condizionale, vale a dire

$$(\mathrm{E}\!\rightarrow) \qquad \frac{A \quad A \rightarrow B}{B}$$

che si chiama *Modus ponens*, in simboli (MP), e che è la regola di eliminazione del condizionale.

Ma nello stesso ragionamento, riformulato nel seguente modo: "Se piove o nevica non esco, dunque se piove non esco"[13], si vede all'opera anche la regola di introduzione del condizionale:

piove

piove o nevica

se piove o nevica non esco

non esco

rappresenta una deduzione di "non esco" da "piove" e "se piove o nevica non esco". Tuttavia è anche accettabile come una deduzione di "se piove non esco" da "se piove o nevica non esco".

La regola, di formulazione un po' più complicata, recita:

(I→) Se assumendo A, oltre ad altre eventuali assunzioni, si deduce B, allora $A \rightarrow B$ si deduce dalle altre eventuali assunzioni, senza A (si dice anche che A viene scaricata[14] dalla deduzione di B nel momento in cui le si aggiunge alla fine $A \rightarrow B$).

La regola qualche volta non è applicata esplicitamente; ci si limita alla deduzione di B da A, come nell'esempio di sopra di "non esco" da "piove". Ma se si vuole riassumere il ragionamento fatto lo si riassume affermando che "quindi $A \rightarrow B$, "quindi se piove non esco".

Si noti che non sempre la parola "allora" indica la presenza di una inferenza. "Piove. Allora non esco" potrebbe essere concepita come una semplice congiunzione. Il ragionamento implicito che fa dipendere la conclusione dalla

[13] "Siccome piove, non esco" non è trasparente, come anche "Non esco perché piove", come è stato discusso in precedenza.

legge "se piove non esco" è ipotetico: chi pronuncia la frase potrebbe non essere in grado di articolare una ragione, ma solo esprimere uno stato d'animo (e in altri momenti uscire anche sotto la pioggia).

Tralasciamo per il momento gli altri connettivi, e vediamo come potremmo dedurre che siccome Giovanni ama Maria allora Giovanni è felice, posto che chi è innamorato di qualcuno è felice. Dobbiamo gestire i quantificatori presenti nelle assunzioni.

Le assunzioni sono che $\forall x(\exists y A(x, y) \to F(x))$ e che $A(g, m)$, dove A è la relazione "essere innamorato di".

Possiamo argomentare come segue:

Giovanni ama Maria

Giovanni è innamorato (di qualcuno)

tutti gli innamorati sono felici

se Giovanni è innamorato è felice

Giovanni è felice

o in formule

$A(g, m)$

$\exists y A(g, y)$

$\forall x(\exists y A(x, y) \to F(x))$

$\exists y A(g, y) \to F(g)$

$F(g)$.

In questa deduzione intervengono due regole dei quantificatori. Una è l'introduzione del quantificatore esistenziale, nel passaggio da $A(g, m)$ a $\exists y A(g, y)$. Benché l'inferenza sembri indebolire il contenuto informativo, essa è di uso frequente e utile, come in questo dialogo:

Giovanni fuma.

Allora c'è ancora qualcuno che fuma.

Per rappresentarla, e per altri usi successivi, indichiamo in generale con $A[x]$, in analogia al caso più semplice di un predicato $P(x)$, un'asserzione in cui occorre la variabile x, priva della determinazione quantitativa data da un quantificatore iniziale. Indichiamo con $A[t]$ il risultato della sostituzione del termine t a tali occorrenze di x.

La regola è:

$$(\mathrm{I}\exists) \qquad \frac{B[t]}{\exists x B[x]}$$

L'altra regola è quella dell'eliminazione del quantificatore universale, nel passaggio da $\forall x(\exists y A(x,y) \to F(x))$ a $\exists y A(g,y) \to F(g)$:

$$(\mathrm{E}\forall) \qquad \frac{\forall x B[x]}{B[t]}$$

Queste due regole non danno problemi e sono ovviamente accettabili nel caso che il termine t sia un nome o una descrizione definita di un individuo, quindi privo di variabili. Nel caso t contenga variabili si rivelano sottili e imprevedibili complicazioni (che coinvolgono lo stesso concetto di sostituzione), che nel linguaggio naturale sono neutralizzate e oscurate dalle flessibilità lessicale, ma su di esse torneremo in seguito. Non ci sono problemi se t coincide con la variabile x: ad esempio da $\forall x \ldots$ si toglie il quantificatore e si continua con "quindi per un x generico ...".

Le ultime due regole, introduzione del quantificatore universale ed eliminazione del quantificatore esistenziale, sono leggermente diverse, in quanto richiedono di prendere in considerazione un ambito non locale, ma la deduzione completa.

Esse tuttavia sono di uso continuo, spesso acritico o inconsapevole, nei ragionamenti.

Supponiamo di voler dimostrare che se un numero x è pari allora $x+1$ è dispari. Si ragiona nel seguente modo: x è pari significa $\exists y(x = 2 \cdot y)$; si dice allora:

$$\text{sia } c \text{ tale che } x = 2 \cdot c.$$

La mossa riassume il ragionamento "introduciamo un nome temporaneo per uno[15] di questi elementi che soddisfano la formula $x = 2 \cdot y$, nome che non può essere uno di quelli disponibili nell'alfabeto perché non si sa quale sia questo elemento[16].

Poi si continua con: allora

$$x + 1 = 2 \cdot c + 1$$

[15] In questo esempio particolare ce ne può essere solo uno, ma in generale, col quantificatore esistenziale, non si sa quanti ≥ 1 ce ne sono.

[16] Siccome c dipende da x, si dovrebbe piuttosto avere una funzione di x, o una notazione del tipo c_x, che però sarebbe inutilmente pesante.

e con altri passaggi, che ora non interessano, ma sui quali torneremo, per concludere che $x + 1$ non è pari.

In questa conclusione, probabilmente[17] della forma $\neg\exists y(x + 1 = 2 \cdot y)$, c non occorre. Quando lo si menziona la prima volta "sia c tale che ..." si presuppone che c sia un simbolo nuovo, che non occorre precedentemente nella deduzione, se c'è una parte precedente, usato solo appunto per eliminare il quantificatore esemplificando a sua affermazione esistenziale, ma senza saper nulla di c se n on che $x = 2 \cdot c$ (e x è generico pure lui). Non ha senso che c compaia nella conclusione della dimostrazione.

La regola di eliminazione del quantificatore esistenziale giustifica tali forme di ragionamento, affermando che

(E$\exists$) Se da $B[c]$ si deduce A, dove c è un nuovo simbolo che non occorre nella parte precedente della deduzione né in A, allora A si deduce da $\exists x B[x]$.

In pratica, quando nel corso di una deduzione si arriva a un $\exists x B[x]$, si aggiunge $B[c]$ e si continua, quando si arriva a una A nella quale c non occorre, si scarica $B[c]$, come se non la si fosse usata, come se A fosse stata dedotta da $\exists x B[x]$ e non da $B[c]$.

Infine la dimostrazione che se x è pari allora $x + 1$ non è pari permette di illustrare l'ultima regola per i quantificatori. Nella dimostrazione si ragiona su x inteso come un numero generico, ma non particolare: si intende che x ha un significato universale. l quantificatore universale non si scrive solo per comodità di espressione, e per lavorare subito algebricamente sulle formule; alla fine si è dedotto "x pari $\rightarrow x + 1$ dispari", e non si è fatto uso di alcun'altra assunzione, oltre alle leggi dell'aritmetica, presupposte (nel corso della deduzione, dovendosi dedurre un condizionale, si è assunto "x pari", ma al fine di invocare (I$\rightarrow$), e poi si è scaricata). Quindi si conclude "$\forall x(x$ pari $\rightarrow x + 1$ dispari)".

(I$\forall$) Se $B[x]$ è dedotto da assunzioni nelle quali non occorre x priva della specificazione quantitativa data da un quantificatore $\forall x$ o $\exists x$, allora si può dedurre $\forall x B[x]$.

Di questa regola non si è normalmente consapevoli perché spesso non se ne fa uso nel linguaggio naturale: si lascia il pronome universale, ad esempio "qualsiasi"; invece nel linguaggio formale occorre fare l'ultimo passaggio, sostituire "qualsiasi" con "uno" e premettere $\forall$.

[17] Diciamo "probabilmente" solo per non entrare in una discussione di come si possono definire "pari" "dispari"; abbiamo usato la definizione proposta in precedenza negli esempi di formalizzazione.

Torniamo ora alla dimostrazione che se x è pari allora $x+1$ è dispari, riprendendola da dove eravamo arrivati, a dedurre cioè $x+1 = 2 \cdot c + 1$.

Per arrivare a $\neg\exists y(x+1 = 2 \cdot y)$, essendo la conclusione desiderata una negazione, si fa una dimostrazione per assurdo. Le dimostrazioni di questo tipo sono codificate dalle regole per la negazione.

(I$\neg$) Se assumendo A, oltre alle altre eventuali assunzioni, si deduce una qualunque contraddizione $B \wedge \neg B$, allora $\neg A$ si deduce dalle altre eventuali assunzioni, senza A (A è scaricata dalla deduzione di $B \wedge \neg B$ nel momento che si aggiunge $\neg A$).

Analogamente

(E$\neg$) Se assumendo $\neg A$, oltre alle altre eventuali assunzioni, si deduce una qualunque c ontraddizione $B \wedge \neg B$, allora A si deduce dalle altre eventuali assunzioni, senza $\neg A$ ($\neg A$ è scaricata dalla deduzione di $B \wedge \neg B$ nel momento che si aggiunge A).

Assumiamo allora (per assurdo, si dice, perché bisogna sapere dove si vuole andare a parare) $\exists y(x+1 = 2 \cdot y)$. Per (E$\exists$), sia d tale che $x+1 = 2 \cdot d$. Ma avevamo già $x+1 = 2 \cdot c+1$, quindi $2 \cdot d = 2 \cdot c+1$. Lasciamo per esercizio di ricavare da qui aritmeticamente una contraddizione[18]. Questo conclude la nostra dimostrazione che se x è pari allora $x+1$ è dispari.

Scriviamo

$$A_1, \ldots, A_n \vdash B$$

per dire che B si deduce dalle assunzioni $A_1, \ldots, A_n$.

B può essere una delle A_i, in particolare $B \vdash B$.

$\vdash B$ significa che esiste una deduzione di B che non fa uso di alcuna assunzione (a parte eventualmente quelle scaricate nel corso della deduzione).

Una deduzione è una successione di frasi ciascuna delle quali si ottiene da precedenti per mezzo di una regola, oppure è una assunzione.

Quando si chiede di dimostrare $A_1, \ldots, A_n \vdash B$ si intende che si deve trovare una deduzione di B che usi le assunzioni $A_1, \ldots, A_n$ (non necessariamente tutte e in questo ordine). Ma è anche possibile dimostrare che esiste una deduzione di un certo tipo se ne esiste un'altra di un altro tipo, senza scriverla esplicitamente; questi risultati diventano tanto più disponibili quanto più si accumulano gli esempi svolti di deduzioni.

[18] Quando si è nel contesto di una teoria particolare, come potrebbe essere l'aritmetica, con contraddizione si intende anche una qualunque formula che ne contraddice altre già dimostrate, ad esempio $0 = 1$.

Ad esempio se si sa che $A_1, \ldots, A_n, A \vdash B$, si può affermare $A_1, \ldots, A_n \vdash A \to B$, senza stare a scrivere la deduzione, che si ottiene da quella di $A_1, \ldots, A_n, A \vdash B$ semplicemente aggiungendo una riga $A \to B$ e scaricando l'assunzione A.

Se (si sa che) $A_1, \ldots, A_n \vdash B$ e nel corso di una deduzione occorrono $A_1, \ldots, A_n$, dedotte non necessariamente in questo ordine o consecutive, si può aggiungere B facendo appello alla deducibilità $A_1, \ldots, A_n \vdash B$. Si evita di inserire nella deduzione in corso la deduzione completa di B da $A_1, \ldots, A_n$, che si potrebbe fare.

Si noti che in una deduzione non tutte le righe e non tutte le assunzioni devono essere utilizzate. Alcune possono fare da spettatrici passive. In particolare si noti che se $A_1, \ldots, A_n \vdash B$ allora $A_1, \ldots, A_n, C \vdash B$.

Bisogna abituarsi a considerare le regole per quello che dicono esattamente, e ad applicarle senza pensare ai ragionamenti reali che le deduzioni devono mimare. Normalmente si evita di enunciare ipotesi inutili[19] ma le regole lo permettono. Bisogna cogliere esattamente tutto e solo quello che le regole permettono.

Due deduzioni si possono quindi mescolare, purché si conservi l'ordine interno alle due successioni, e si ottiene ancora una deduzione, di quella che risulta la sua ultima riga.

Quando si sfruttano in questo modo precedenti risultati di deducibilità per abbreviare nuove deduzioni si dice che si fa uso di *regole derivate*.

Esempio 1.2.1. $A \to B, \neg B \vdash \neg A$.

1	A	assunzione da scaricare al passo 6
2	$A \to B$	assunzione
3	B	(MP) da 1 e 2
4	$\neg B$	assunzione
5	$B \wedge \neg B$	(I$\wedge$) da 3 e 4
6	$\neg A$	(I$\neg$) da 1 e 5

Si dice che $\neg A$ è ottenuta per contrapposizione da $A \to B$ e $\neg B$.

Ne segue anche la seguente forma della contrapposizione:

$$A \to B \vdash \neg B \to \neg A.$$

Naturalmente quando si inizia una derivazione come quella di sopra, non si sa ancora quando verrà scaricata la prima assunzione; quindi in pratica la

[19] Ma se uno studente si presenta all'iscrizione con più documenti di quelli richiesti, non per questo l'iscrizione gli è rifiutata.

derivazione si presenta nel seguente modo:

1	A	assunzione da scaricare
2	$A \to B$	assunzione
3	B	(MP) da 1 e 2
4	$\neg B$	assunzione
5	$B \wedge \neg B$	(I$\wedge$) da 3 e 4
6	$\neg A$	(I$\neg$) da 1 e 5

e quando si arriva al passo 6 e si vede che si può scaricare la 1 si può scrivere

1	A	assunzione da scaricare al passo 6
2	$A \to B$	assunzione
3	B	(MP) da 1 e 2
4	$\neg B$	assunzione
5	$B \wedge \neg B$	(I$\wedge$) da 3 e 4
6	$\neg A$	(I$\neg$) da 1 e 5

cancellando esplicitamente, per chiarezza, l'assunzione scaricata (o meglio la sua caratterizzazione come assunzione da scaricare).

Esempio 1.2.2. $A \to B, \neg A \to B \vdash B$.

1	$\neg B$	assunzione da scaricare al passo 8
2	$A \to B$	assunzione
3	$\neg A$	contrapposizione da 1 e 2
4	$\neg A \to B$	assunzione
5	$\neg B \to \neg\neg A$	contrapposizione da 5
6	$\neg\neg A$	(MP) da 1 e 5
7	$\neg A \wedge \neg\neg A$	(I$\wedge$) da 3 e 6
8	B	(E$\neg$) da 1 e 7

oppure

1	$\neg B$	assunzione da scaricare al passo 7
2	$A \to B$	assunzione
3	$\neg A$	contrapposizione da 1 e 2
4	$\neg A \to B$	assunzione
5	B	(MP) da 3 e 4
6	$B \wedge \neg B$	(I$\wedge$) da 5 e 1
7	B	(E$\neg$) da 1 e 6

Con passaggi simili a quelli dell'ultimo esempio si fa vedere che se

$$A_1, \ldots, A_n, B \vdash C$$

e

$$A_1, \ldots, A_n, \neg B \vdash C,$$

allora

$$A_1, \ldots, A_n \vdash C.$$

(Dal *merging* delle due deduzioni si ottiene sia $B \to C$ sia $\neg B \to C$, dedotte da $A_1, \ldots, A_n$, quindi sia $\neg C \to \neg B$ sia $\neg C \to \neg\neg B$; quindi se si assume $\neg C$ si ottiene una contraddizione, quindi C.)

In questo modo si dimostra l'importante, e utile

$$A \to C, B \to C \vdash A \vee B \to C,$$

detta "dimostrazione per casi", o "per distinzione di casi". Basta dimostrare

$$A \to C, B \to C, A \vee B, B \vdash C,$$

con

1	B	assunzione
2	$B \to C$	assunzione
3	C	(MP) da 1e 2

e

$$A \to C, B \to C, A \vee B, \neg B \vdash C,$$

con

1	$\neg B$	assunzione
2	$A \vee B$	assunzione
3	A	(E$\neg$) da 1 e 2
4	$A \to C$	assunzione
5	C	(MP) da 3 e 4.

La contrapposizione e la dimostrazione per casi sono esempi di possibili regole derivate.

Allo stesso modo si dimostra:

$$\vdash A \vee \neg A.$$

Viceversa se si ammette la regola della distinzione di casi e $A \vee \neg A$ si dimostra

$$A \to B, \neg A \to B \vdash B$$

e con questa

$$\neg\neg A \vdash A$$

(si veda l'Esercizio 1.2.3 parti 5, 6 e 7 qui sotto) che rende superflua la regola (E$\neg$).

1.2.1 Esercizi

Esercizio 1.2.3. Dimostrare le seguenti:

1. $A \to B \vdash \neg A \vee B$.
2. $\neg A \vee B \vdash A \to B$.
3. $\neg A \vee \neg B \vdash \neg(A \wedge B)$.
4. $\neg(A \wedge B) \vdash \neg A \vee \neg B$.
5. $\vdash A \to (\neg A \to B)$.
6. $A, \neg A \vdash B$.
7. $\vdash A \to (B \to A)$.
8. $\vdash \neg\neg A \to A$.
9. $\vdash (\neg A \to A) \to A$.
10. $\neg A \vee \neg B \vdash \neg(A \wedge B)$.
11. Dimostrare che se $\vdash A \to B$ e $\vdash B \to C$ allora $\vdash A \to C$.

    Più in generale, se

    $$A_1, \ldots, A_n \vdash A \quad \text{e} \quad B_1 \ldots, B_m, A \vdash B$$

    allora $A_1, \ldots, A_n, B_1, \ldots, B_m \vdash B$.

Esercizio 1.2.4. Trovare esempi di ragionamento in (buon) linguaggio naturale o matematico che siano formalizzabili da

$$\neg A \vee \neg B \vdash \neg(A \wedge B),$$
$$A \to (B \to C), A \to B \vdash A \to C.$$

Esercizio 1.2.5. Trovare deduzioni che mostrino che i seguenti ragionamenti sono corretti:

1. I cavalli sono quadrupedi. Qualche cavallo è veloce. Quindi qualche quadrupede è veloce.
2. Tutte le rose hanno le spine. Le spine fanno male. Dunque le rose fanno male.

3. Nessun rivoluzionario è moderato. Ma qualche progressista è moderato. Quindi ci sono progressisti che non sono rivoluzionari.

Esercizio 1.2.6. Riprendere l'Esercizio 1.1.1 a pagina 8, e visto che si tratta di ragionamenti, controllare se la conclusione è deducibile dalla premessa, cioè se

$$\neg\mathrm{Prev} \to (\neg\mathrm{Det} \vee \neg\mathrm{Raz}) \vdash \mathrm{Det} \to (\mathrm{Prev} \to \mathrm{Raz})$$

e se

$$\neg\mathrm{Prev} \to (\neg\mathrm{Det} \vee \neg\mathrm{Raz}) \vdash \mathrm{Det} \to (\neg\mathrm{Prev} \to \neg\mathrm{Raz}).$$

1.2.2 Appendice: sulla logica intuizionistica

I sistemi di regole deduttive, come quello che abbiamo presentato, sono abbastanza arbitrari. Il formato delle deduzioni, la scelta delle regole di base rispetto a quelle derivate dipendono dall'interesse che guida la costruzione del sistema, che può andare dalla ricerca della efficienza o semplicità a quella dell'analisi filosofica.

In ogni caso di solito un requisito che si chiede è che le regole siano complete, nel senso di permettere di formalizzare tutti i ragionamenti validi in un determinato dominio di conoscenza.

L'argomento sarà approfondito a tempo debito. Per ora facciamo vedere come si potrebbe presentare un sistema equivalente di regole di deduzione naturale, per quel che riguarda i connettivi.

Invece della regola (E∨) vista sopra, che coinvolge la negazione oltre alla disgiunzione, e si può perciò considerare non pura, si può assumere come regola di base per la disgiunzione la distinzione dei casi, schematicamente

$$\frac{A \vee B \quad \begin{array}{c}\overline{A}\\ \vdots\\ C\end{array} \quad \begin{array}{c}\overline{B}\\ \vdots\\ C\end{array}}{C}$$

con lo scarico di A e di B, indicato dalla sopralineatura.

Per quel che riguarda la negazione, si mantenga la (I¬) e come eliminazione della negazione di consideri

$$\frac{A \quad \neg A}{C}$$

per una C qualunque.

Si tratta della legge detta *ex falso quodlibet*, dove "falsum" significa "contraddittorio" (e "quodlibet" significa "qualunque cosa"). La sua naturalità è confermata da affermazioni come la seguente, che capita di sentire (si veda anche esercizio 1.2.3 parte 6):

> Prima dici una cosa e poi dici il contrario, sei inaffidabile: allora da te ci si può aspettare qualunque cosa.

Scriviamo $A_1, \ldots, A_n \vdash_I B$ se esiste una deduzione di B da $A_1, \ldots, A_n$, dove I per "intuizionistico".

Il sistema di regole per cui la deducibilità è stata indicata con $\vdash$ si chiama invece sistema classico.

Esempi

Esempio 1.2.7. $A \vdash_I \neg\neg A$, con la deduzione

1	A	assunzione
2	$\neg A$	assunzione da scaricare al passo 3
3	$\neg\neg A$	*ex falso* $\ldots \neg\neg A$.

Esempio 1.2.8. $\vdash_I \neg\neg(A \vee \neg A)$. Questa seguirebbe con la deduzione

1	$\neg(A \vee \neg A)$	assunzione da scaricare al passo 5
2	$\neg A \wedge \neg\neg A$	??
3	$\neg A$	(E$\wedge$)
4	$\neg\neg A$	(E$\wedge$)
5	$\neg\neg(A \vee \neg A)$	*ex falso* $\ldots$

e il passo 2 è in effetti giustificato dal seguente

Esempio 1.2.9. $\neg(A \vee B) \vdash_I \neg A \wedge \neg B$ con la deduzione

1	$\neg(A \vee B)$	assunzione
2	A	assunzione da scaricare al passo 4
3	$A \vee B$	(I$\vee$)
4	$\neg A$	(I$\neg$)
5	B	assunzione da scaricare al passo 7
6	$A \vee B$	(I$\vee$)
7	$\neg B$	(I$\neg$)
8	$\neg A \wedge \neg B$	(I$\wedge$), da 4 e 7.

Non si riesce invece a ottenere

$$\vdash_I A \vee \neg A$$

Se si avesse questa deduzione, si potrebbe dimostrare come si è visto

$$\neg\neg A \vdash_I A$$

ma una deduzione di tale legge, che è equivalente alla vecchia regola (E$\neg$) (esercizio), non esiste nel sistema I.

Come si possa *dimostrare* in generale che non esiste una deduzione in un dato sistema lo discuteremo più avanti.

La vecchia regola di eliminazione della disgiunzione è invece ora ottenibile come regola derivata, in quanto

Esempio 1.2.10. $\neg A, A \vee B \vdash_I B$.

Infatti se oltre a $\neg A$ si assume A, si ha B per *ex falso quodlibet*; se oltre a $\neg A$ si assume B si ha B, banalmente, quindi B segue da $\neg A$ e $A \vee B$.

Se quindi si aggiunge alle regole del sistema I la regola

$$\frac{\neg\neg A}{A}$$

si ottiene un sistema equivalente a quello classico, per quanto riguarda i connettivi.

Capitolo 2

Logica proposizionale

La logica proposizionale studia quello che si può dire sulle frasi considerando solo la struttura determinata dai connettivi; si considerano cioè frasi formate da frasi più semplici per mezzo dei connettivi.

Per fare questo la logica matematica costruisce un preciso modello matematico i cui oggetti sono la versione formale delle frasi composte con i connettivi.

Introdotti i simboli per i connettivi, occorre dare le loro precise regole d'uso sia dal punto di vista sintattico (dove scriviamo ad esempio $\neg$ per formare la negazione di un'asserzione?), sia da quello deduttivo (già anticipato), sia da quello semantico (come interpretiamo il significato delle frasi composte, in funzione delle frasi componenti?).

2.1 Sintassi

La necessità di fornire regole rigide per la formazione delle frasi è data dalla volontà di evitare le ambiguità possibili nelle lingue naturali. Alcune ambiguità si riferiscono proprio alla distribuzione dei connettivi; supponiamo ad esempio di leggere un problema rappresentato dall'insieme delle seguenti disequazioni:

$$x^2 + 4x + 3 < 0 \quad \text{e} \quad x < -3 \quad \text{o} \quad x > -2.$$

Uno tende a rispondere "risolvo la prima, poi interseco con $x < -3$ e unisco con $x > -2$", ma è proprio l'ordine di queste operazioni che conta, che non sempre è quello del *first come, first served*, con la scrittura da sinistra a destra.

Si può intendere che si chieda quali siano i valori per cui si ha che

$$x^2 + 4x + 3 < 0 \quad \text{e} \quad x < -3$$

oppure si ha che

$$x > -2;$$

si può anche intendere che si chieda quali siano i valori per cui si ha

$$x^2 + 4x + 3 < 0$$

ma ristretti ad essere

$$x < -3 \quad \text{o} \quad x > -2.$$

Nel primo caso la risposta è $(-2, +\infty)$, nel secondo caso è $(-2, -1)$.

Naturalmente l'ambiguità, che nel parlato si risolve con le pause, nella scrittura matematica si risolve con le parentesi, il primo caso essendo

$$(x^2 + 4x + 3 < 0 \quad \text{e} \quad x < -3) \quad \text{o} \quad x > -2$$

e il secondo caso

$$x^2 + 4x + 3 < 0 \quad \text{e} \quad (x < -3 \quad \text{o} \quad x > -2).$$

La stessa soluzione delle parentesi[1] adotteremo per le formule logiche.

Il linguaggio proposizionale

Le frasi di ogni linguaggio sono stringhe[2] di simboli dell'alfabeto. L'alfabeto del *linguaggio proposizionale* contiene, oltre ai connettivi, le parentesi sinistra "(" e destra ")", e un insieme $\mathcal{L}$ di lettere, dette lettere proposizionali. Lo supponiamo infinito per averne sempre a disposizione quante ne servono.

Tali lettere si chiamiamo anche variabili proposizionali, ma preferiamo non seguire questo uso perché il loro dominio di variabilità (le frasi) è per ora troppo indefinito[3].

Le parole accettabili di questo alfabeto si chiameranno *proposizioni*, o anche formule, che sono termini tecnici per distinguerle dalle asserzioni dei linguaggi dotati di senso. Quello che importa delle proposizioni è solo la loro struttura formale, che poi si dovrà riconoscere nelle frasi dei linguaggi naturali o matematici, quando il linguaggio proposizionale sarà interpretato sostituendo alle lettere frasi relative ad un determinato argomento.

[1]Le parentesi sono state anche aggiunte al linguaggio naturale — almeno nella saggistica, meno in letteratura — con un'altra funzione, quella di racchiudere un inciso (una frase parentetica appunto) non di articolare una frase complessa.

[2]Con "stringa" s'intende una lista o una successione finita. Non è necessario entrare nei particolari del tipo di rappresentazione dei dati che si sceglie, finché non si deve implementare.

[3]Quando introdurremo la semantica formale diventerà preciso.

Non tutte le stringhe di simboli dell'alfabeto sono ammesse come proposizioni. Una generica stringa, anche illecita, è chiamata "parola".

La definizione dell'insieme $\mathcal{P}$ delle proposizioni stipula innanzi tutto che:

Per ogni lettera $p \in \mathcal{L}$, (p) è una proposizione, detta *atomica*

(cioè non ulteriormente analizzata e scomposta nel contesto della trattazione).

Per il resto della definizione, occorre parlare di proposizioni qualunque e della loro composizione; è quindi necessario avere delle variabili che variano sull'insieme delle proposizioni, e che si chiamano *metavariabili*[4]; useremo le lettere $A, B, \ldots$

Si danno quindi le seguenti clausole:

1. Se A è una proposizione, anche $(\neg A)$ lo è.
2. Se $\bullet$ è un connettivo binario, e se A e B sono proposizioni, anche $(A \bullet B)$ lo è.

Le clausole della definizione sono anche regole di costruzione. S'intende che ogni proposizione si ottiene applicando un numero finito di volte le clausole della definizione.

Esempi 2.1.1. 1. $(p \cap q)$ non è una proposizione perché $\cap$ non è un elemento dell'alfabeto[5].

2. $p \wedge q$ non è una proposizione, perché:

   Ogni proposizione contiene almeno una parentesi[6]

3. $)p($ non è una proposizione, come non lo sono p o $)p)$, perché:

   Ogni proposizione inizia con una parentesi (e termina con una parentesi).

[4]La ragione di questo termine, non usato altrove in matematica, se non in logica, è che queste variabili indicano elementi di una struttura che è anch'essa un linguaggio, e che contiene a sua volta variabili (le lettere) che devono essere interpretate su frasi; "meta" significa "sopra", "oltre", e deriva dal greco, dove significava piuttosto "dopo"; ma da che sono stati chiamati "metafisica" i libri di Aristotele che seguivano quelli di fisca, è venuta questa variante di significato.

[5]Per i linguaggi formali si chiede sempre che l'alfabeto e le sue diverse categorie di simboli siano insiemi *decidibili*, cioè tali che l'appartenenza o meno ad essi di un simbolo possa essere decisa da un algoritmo.

[6]Queste e le proprietà si dimostrano rigorosamente per induzione sulla altezza delle proposizioni, basandosi sulla definizione delle stesse, che è una definizione induttiva. Si veda più avanti la definizione di "altezza" di una proposizione.

4. $((p) \to (q))$ è una proposizione perché ottenuta dalle proposizioni atomiche (p) e (q) con una applicazione della clausola induttiva relativa a $\to$.

5. $(\neg((p) \to (q)))$ è una proposizione perché ottenuta dalle proposizioni atomiche (p) e (q) con una prima applicazione della clausola induttiva relativa a $\to$ e una seconda applicazione della clausola relativa a $\neg$.

6. $((p)$ non è una proposizione perché:

   In ogni proposizione il numero di parentesi sinistre è uguale al numero di parentesi destre.

7. (pq) non è una proposizione perché non è atomica e non contiene nessun connettivo.

Se una proposizione è della forma $(\neg A)$ o della forma $(A \bullet B)$, $\neg$ e $\bullet$ sono rispettivamente il suo connettivo principale, e A e B le sottoproposizioni immediate.

Si dice che $(\neg A)$ è una negazione, citando il suo connettivo principale, la negazione di A — e si legge "non A"; si dice che $(A \wedge B)$ è una congiunzione, la congiunzione di A e B — e si legge "A e B"; A e B sono le proposizioni congiunte in $(A \wedge B)$; analogamente per la disgiunzione[7] $(A \vee B)$ — che si legge "A o B"; $(A \oplus B)$ si può leggere "o A o B"; $(A \to B)$ si dice un condizionale — e si legge di solito "se A allora B"; A si chiama antecedente, e B conseguente; $(A \leftrightarrow B)$ si dice bicondizionale — e si legge "A se e solo se B".

Analisi sintattica

Una proposizione è una lista di simboli, ma è anche passibile di una rappresentazione con una diversa struttura. A ogni proposizione è associato un *albero di costruzione*, o di *analisi sintattica*[8], che è un albero etichettato finito binario.

Un albero binario[9] è un insieme X parzialmente ordinato, cioè con una relazione $\preceq$ con le seguenti proprietà: $\preceq$ è una relazione riflessiva, transitiva e antisimmetrica[10]. Gli elementi dell'albero si chiamano *nodi*. Se $x \preceq y$, si dice che y è un successore, o un discendente di x. Esiste un nodo minimo

[7] Chiameremo $\vee$ semplicemente disgiunzione, e $\oplus$ disgiunzione esclusiva o forte.

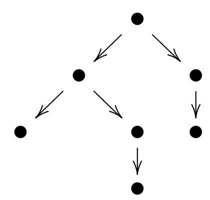
[8] In inglese *parsing*.

[9] Esistono definizioni leggermente diverse, più o meno generali, ad esempio con una o più radici; diamo quella che serve ai nostri scopi.

[10] Questo significa che: (1) $x \preceq x$, (2) se $x \preceq y$ e $y \preceq z$ allora $x \preceq z$ e (3) se $x \preceq y$ e $y \preceq x$ allora $x = y$.

r tale che $r \preceq x$ per ogni nodo di X, e si chiama *radice*. I nodi a tali che non esiste $b \neq a$ per cui $a \preceq b$ si chiamano foglie[11]. Ogni nodo che non sia una foglia ha uno o al massimo due successori immediati[12], dove si dice che b è un successore immediato di a se $a \preceq b$, $a \neq b$ e non esiste un c tale che $a \preceq c \preceq b$, con $c \neq a$ e $c \neq b$.

La rappresentazione usuale di un albero binario è di questo tipo:



dove con la freccia si indica il successore immediato, la radice è in alto e l'albero cresce verso il basso.

Un *ramo* è un insieme totalmente ordinato[13] di nodi che va dalla radice a una foglia. La sua lunghezza è il numero di nodi che vi appartengono. L'altezza dell'albero è la massima lunghezza dei suoi nodi.

Un albero si dice etichettato se ad ogni nodo è associato un elemento di qualche insieme prefissato, che si chiama etichetta (*label*). Le etichette si possono sovrapporre ed identificare con i nodi.

L'albero sintattico di una proposizione è definito in questo modo:

- la radice è (etichettata con) la proposizione data

- ogni nodo ha nessuno, uno o due successori immediati a seconda che la proposizione etichetta del nodo sia atomica, o della forma $(\neg A)$, o della forma $(A \bullet B)$. Nel secondo caso il successore è etichettato con A, nel terzo caso i due successori sono etichettati rispettivamente con A e con B.

Si chiama *altezza* della proposizione l'altezza del suo albero di costruzione.

[11] Esistono sempre se l'albero, ovvero l'insieme dei nodi X, è finito.

[12] Un'altra terminologia è "figli". Se ci sono due figli, s'intende che sono esplicitamente distinti il primo e il secondo – sulla pagina, a sinistra e a destra.

[13] Per ora basti intendere che due nodi qualunque del ramo sono confrontabili, quindi ogni nodo del ramo salvo l'ultimo ha esattamente un successore immediato, e con ogni nodo nel ramo ci sono tutti i suoi precedenti.

Esempio 2.1.2. L'albero per $(((p) \wedge (\neg(q))) \vee (\neg(p)))$ è il seguente:

$$
\begin{tikzcd}
 & & & (((p) \wedge (\neg(q))) \vee (\neg(p))) \arrow[dll] \arrow[dr] & \\
 & ((p) \wedge (\neg(q))) \arrow[dl] \arrow[dr] & & & (\neg(p)) \arrow[d] \\
(p) & & (\neg(q)) \arrow[d] & & (p) \\
 & & (q) & &
\end{tikzcd}
$$

La sua altezza è quattro.

Le etichette dei nodi dell'albero di costruzione di una proposizione sono le sue *sottoproposizioni*. Le lettere che compaiono nelle (proposizioni atomiche nelle) foglie sono le lettere che occorrono nella proposizione; si dice che un simbolo occorre in una proposizione se è un elemento della lista (che è la proposizione); le occorrenze di un simbolo in una proposizione sono i vari posti della lista in cui il simbolo si presenta. Se $p, \ldots, q$ sono le lettere che occorrono nella proposizione A, si scrive anche $A[p, \ldots, q]$. Qualche volta si usa questa notazione anche se $p, \ldots, q$ sono solo alcune delle lettere che occorrono in A, o viceversa se le lettere che occorrono in A sono incluse tra le $p, \ldots, q$; invece di introdurre notazioni distinte apposite, la differenza sarà chiarita dal contesto o da esplicite precisazioni.

Le parentesi sono essenziali per individuare il connettivo principale di una proposizione, e quindi per costruire il suo albero sintattico. Per individuare il connettivo principale, si usa un contatore di parentesi[14].

Il contatore scansisce la lista da sinistra verso destra, e scatta di $+1$ quando incontra una parentesi sinistra, di -1 quando incontra una parentesi destra. Condizione necessaria affinché una parola sia una proposizione è che il contatore, inizializzato a 0, non assuma mai valori negativi e torni a 0 solo alla fine della parola. Perché poi la parola sia una proposizione bisogna che gli altri simboli siano distribuiti in mezzo alle parentesi in modo corretto.

Il fatto che in una proposizione il contatore alla fine vale 0 dipende dalla proprietà che in una proposizione il numero di parentesi sinistre è uguale a quello delle parentesi destre. Che il contatore non possa tornare a 0 prima della fine dipende invece dalla proprietà che in ogni proposizione, se si considera un suo segmento iniziale proprio, nel segmento il numero di parentesi

[14]Horstmann, p. 76.

sinistre è maggiore di quello delle parentesi destre. Questo è anche il motivo per cui il contatore non può assumere valori negativi.

Ad esempio per

$$(((p) \land (\neg(q))) \lor (\neg(p)))$$

il contatore assume i valori:

$$1\ 2\ 3\ 2\ 3\ 4\ 3\ 2\ 1\ 2\ 3\ 2\ 1\ 0.$$

Una prima condizione necessaria è soddisfatta.

Per individuare il suo possibile connettivo principale, si elimina la coppia di parentesi esterne, e si mette di nuovo in funzione il contatore su

$$((p) \land (\neg(q))) \lor (\neg(p))$$

notando che esso assume questa volta i valori

$$1\ 2\ 1\ 2\ 3\ 2\ 1\ 0\ 1\ 2\ 1\ 0$$

tornando a 0 la prima volta[15] quando arriva alla fine di

$$((p) \land (\neg(q))),$$

e a destra c'è un connettivo binario.

La parola $((p) \land (\neg(q)))$ è candidata ad essere una sottoproposizione; il connettivo $\lor$ che compare nel prossimo posto a destra è candidato a essere il connettivo principale; l'ipotesi cioè è che se il procedimento va a buon fine la proposizione sia della forma $(A \lor B)$. Se anche la parte destra restante

$$(\neg(p))$$

è una proposizione si sarà trovato che la parola data è la disgiunzione di $((p) \land (\neg(q))$ e di $(\neg(p))$.

Poiché quest'ultima si vede facilmente che è una proposizione, proseguiamo l'analisi di

$$((p) \land (\neg(q))).$$

Il contatore applicato a questa assume i valori

$$1\ 2\ 1\ 2\ 3\ 2\ 1\ 0$$

e applicato a

$$(p) \land (\neg(q))$$

[15]Invece di eliminare le parentesi esterne e di prendere nota di dove il contatore torna a 0 sulla parola ridotta, si può ovviamente prendere nota di dove torna a 1 in quella originaria.

i valori

$$1\ 0\ 1\ 2\ 1\ 0$$

tornando la prima volta a 0 alla fine di (p), individuando a destra il connettivo $\wedge$, che lega (p) e $(\neg(q))$. In questo modo si arriva a costruire l'albero sintattico.

Invece

$$(((p) \wedge (q(\neg))) \vee (\neg(p)))$$

non è una proposizione, nonostante il contatore si comporti come nel caso precedente (esercizio).

Oltre ai casi che si presentano nel precedente esempio, un altro possibile è il seguente. Data la parola

$$(\neg((p) \to ((p) \oplus (q))))$$

il contatore assume i valori

$$1\ 2\ 3\ 2\ 3\ 4\ 3\ 4\ 3\ 2\ 1\ 0$$

ma se si tolgono le parentesi esterne e si riapplica il contatore, esso torna a 0 solo alla fine. In questo caso il connettivo, che deve esserci perché è ovvio che non si tratta di formula atomica, non può che essere la negazione, all'inizio. Infatti in questo caso c'è. Se non ci fosse, la parola sarebbe della forma $((A))$, o $(\bullet \ldots$, che non sono proposizioni.

Alcune parentesi sono sovrabbondanti, ma solo quelle della coppia più esterna e quelle nelle proposizioni atomiche, dove sono usate sia per uniformità sia per sottolineare la differenza tra una lettera come elemento dell'alfabeto e la lettera come proposizione[16]. Ma ora per comodità di scrittura e lettura è meglio ridurre il numero di parentesi con le seguenti convenzioni: non si scrivono le parentesi intorno alle lettere nelle proposizioni atomiche, non si scrivono le parentesi più esterne, e si eliminano alcune coppie di parentesi intorno ad alcune sottoproposizioni, con un criterio sufficiente a farle ripristinare in modo corretto e univoco che è formulato nel seguente modo.

Si ordinano per priorità i connettivi secondo le seguente graduatoria:

$$\begin{array}{c} \neg \\ \wedge \\ \vee \\ \oplus \\ \to \\ \leftrightarrow \end{array}$$

[16] Tra simboli dell'alfabeto e parole c'è una differenza di tipo logico. Nei linguaggi naturali si presentano alcune eccezioni, ma solo le vocali "a", "e", "i", "o" sono usate come parole; tuttavia è raro che si parli dell'alfabeto; quando lo si fa, si scrive appunto "e" e non e.

Data quindi una parola le cui parentesi non rispettano le condizioni per essere una proposizione (sì però la parità, il fatto che il numero di parentesi sinistre sia uguale a quello delle parentesi destre, il fatto che in ogni punto che non sia l'ultimo il numero di sinistre è maggiore o uguale di quello delle destre, e tutte le proprietà che si mantengono quando si eliminano alcune *coppie* di parentesi corrispondenti) le parentesi si rimettono secondo questo procedimento: prima si rimettono le parentesi a sinistra e a destra delle lettere; quindi si prende in esame la negazione, se occorre nella parola; si esamina un'occorrenza della negazione che non abbia immediatamente alla sua destra un'altra negazione[17]. Alla sua destra c'è una parentesi sinistra — altrimenti si può dire che quella parola non proviene dalla eliminazione di coppie di parentesi da una genuina proposizione (brevemente, che non è una proposizione). Sia σ la parola alla sua destra che termina con la parentesi destra che chiude la parentesi sinistra. Per trovare la parentesi destra che "chiude" la parentesi sinistra si usa di nuovo il contatore in modo ovvio. Allora si rimette una parentesi sinistra alla sinistra della negazione, se non c'è già, e una parentesi destra a destra di σ, se non c'è già, ottenendo $(\neg\sigma)$; si ripete per ogni occorrenza di $\neg$, quindi si passa ai connettivi binari. Per ciascuno di essi, indicato con $\bullet$, nell'ordine di priorità, si considerano le più corte sottoparole σ e τ a sinistra e a destra di $\bullet$ che sono chiuse tra due parentesi sinistre e destre, e si introduce una parentesi (a sinistra di σ e) a destra di τ, se non ci sono già, ottenendo $(\sigma \bullet \tau)$, e così via[18].

Per occorrenze multiple dello stesso connettivo si prende in esame[19] l'ultima, quella più a destra; questo significa che per formule composte con uno stesso connettivo ripetuto si conviene l'associazione a destra, cioè ad esempio con $A \to B \to C$ si intende $A \to (B \to C)$, e con $A \wedge B \wedge C$ si intende $A \wedge (B \wedge C)$.

Esempio 2.1.3. Data $p \wedge \neg q \vee \neg p$, la reintroduzione delle parentesi avviene attraverso questa successione di passi:

1 $\quad (p) \wedge \neg(q) \vee \neg(p)$

[17] A parte questa condizione, l'ordine in cui si lavora sulle eventuali diverse occorrenze della negazione, se ce ne è più di una, non è rilevante; lo si può anche (immaginare di) fare in simultanea. Un calcolatore lo può fare in parallelo. Lo stesso vale per gli altri connettivi. Per coerenza con il caso della negazione, per ogni connettivo prenderemo in esame prima la sua ultima (più a destra) occorrenza.

[18] La reintroduzione delle parentesi intorno alle lettere praticamente si può fare anche alla fine, per non appesantire la scrittura, ma si faccia attenzione che in questo caso, nel procedimento sotto descritto, adiacenti ai connettivi si possono trovare anche lettere, oltre a parole che sono delimitate da parentesi. In particolare la negazione che non abbia alla sua destra un'altra negazione può avere o una parentesi, e si procede come nel testo, oppure una lettera p e allora si introducono le parentesi $(\neg p)$. Si veda qui sotto l'esempio.

[19] Come è stato detto sopra nella nota 17.

2 $\quad (p) \land \neg(q) \lor (\neg(p))$

3 $\quad (p) \land (\neg(q)) \lor (\neg(p))$

4 $\quad ((p) \land (\neg(q))) \lor (\neg(p))$

5 $\quad (((p) \land (\neg(q))) \lor (\neg(p)))$.

I passi 2 e 3 si possono naturalmente fare in parallelo.

Data $p \to \neg(q \land \neg\neg r)$

1 $\quad (p) \to \neg((q) \land \neg\neg(r))$

2 $\quad (p) \to \neg((q) \land \neg(\neg(r)))$

3 $\quad (p) \to \neg((q) \land (\neg(\neg(r))))$

4 $\quad (p) \to (\neg((q) \land (\neg(\neg(r)))))$

5 $\quad ((p) \to (\neg((q) \land (\neg(\neg(r))))))$

oppure, per rendere più chiara la lettura

1 $\quad p \to \neg(q \land \neg(\neg r))$

2 $\quad p \to \neg(q \land (\neg(\neg r)))$

3 $\quad p \to (\neg(q \land (\neg(\neg r))))$

4 $\quad (p \to (\neg(q \land (\neg(\neg r)))))$

rimettendo infine le parentesi intorno alle lettere.

Si noti che se fosse stata data $p \to \neg q \land \neg\neg r$ la reintroduzione delle parentesi avrebbe portato a una diversa proposizione:

$$((p) \to ((\neg(q)) \land (\neg(\neg(r)))))$$

(esercizio, e si confrontino i due alberi sintattici), per cui le due parentesi lasciate in $p \to \neg(q \land \neg\neg r)$ sono essenziali, se si vuole parlare della proposizione $((p) \to (\neg((q) \land (\neg(\neg(r))))))$.

Non è comunque necessario né obbligatorio togliere tutte le parentesi; per agevolare la lettura, o all'inizio quando non si è ancora fatta esperienza, può essere conveniente lasciarne alcune, che pure grazie alle convenzioni si potrebbero eliminare. Così ad esempio si potrà scrivere $p \to (q \land r)$ invece di $p \to q \land r$ oppure $(p \lor q) \to r$ invece di $p \lor q \to r$.

Le parentesi si rimettono solo se si ha necessità di capire quale è il connettivo principale, per svolgere l'analisi sintattica. Le parentesi esterne possono tranquillamente essere tralasciate, finché la proposizione non deve essere combinata con altre mediante qualche connettivo — allora si devono rimettere.

L'albero sintattico si può costruire direttamente anche per le espressioni prive di tutte le parentesi, se si tiene presente la priorità dei connettivi. Il connettivo principale è sempre quello di priorità più bassa.

Esempio 2.1.4. L'albero sintattico per $p \land \neg q \lor \neg p$ è il seguente, essendo $\lor$ il connettivo principale:

$$
\begin{array}{cccc}
 & & p \land \neg q \lor \neg p & \\
 & & \swarrow \quad \searrow & \\
 & p \land \neg q & & \neg p \\
 \swarrow & \searrow & & \downarrow \\
 p & \neg q & & p \\
 & \downarrow & & \\
 & q. & &
\end{array}
$$

Le etichette sono diverse, ma l'albero è lo stesso della proposizione analizzata in precedenza.

Dalla prossima sezione 2.2, chiameremo "proposizioni" anche le parole ottenute da proposizioni per eliminazione di parentesi rispettando la convenzione sulla priorità.

Esercizi

Esercizio 2.1.5. Discutere se le seguenti parole sono proposizioni:

$$
\begin{array}{c}
(p \land (q) \\
(p)) \land q) \\
((p) \land q) \\
((p) \land (\neg(q))) \\
((p) \to \land) \\
p \\
((p))
\end{array}
$$

Esercizio 2.1.6. Verificare quali delle seguenti parole sono proposizioni — secondo la definizione originaria — e quali no, costruendo l'albero sintattico e spiegando dove eventualmente la costruzione fallisce e per quale ragione:

$$
\begin{gathered}
(\neg(\neg p)) \\
((p) \to ((q) \vee (\neg(r)))) \\
(\neg\neg((p) \to (q))) \\
((((p) \to (q)) \wedge (p)) \to (q)) \\
((\neg(p)) \wedge (q)) \vee (r)) \\
(((\neg(p)) \wedge (q)) \vee (r)) \\
((p) \wedge (q) \wedge (r)).
\end{gathered}
$$

Esercizio 2.1.7. Dare ragioni per le seguenti proprietà:

- Ogni proposizione ha lunghezza maggiore o uguale a 3.
- In ogni proposizione non atomica occorre un connettivo.
- In nessuna proposizione occorrono due connettivi consecutivi.
- In nessuna proposizione occorre la sottosequenza (), né $)p$.
- In ogni proposizione la sua lunghezza (come lista) è maggiore della sua altezza.
- In ogni proposizione, ogni suo segmento iniziale proprio contiene più parentesi sinistre che destre.

  Suggerimento: la dimostrazione di queste proprietà è per induzione sulla altezza delle proposizioni: si dimostrano prima per le proposizioni (p), quindi supponendo che valgano per proposizioni A, B si dimostra che valgono anche per $(\neg A)$ e $(A \bullet B)$.

Esercizio 2.1.8. Una misura di complessità delle proposizioni è una funzione dalle proposizioni nei numeri naturali che soddisfa la condizione che la misura di una proposizione è maggiore delle misure delle proposizioni componenti, e le atomiche hanno tutte la stessa misura minima. Il numero (di occorrenze) dei connettivi è una misura di complessità, come lo sono la lunghezza (della stringa) e l'altezza (dell'albero sintattico).

Trovare la relazione tra il numero di occorrenze di connettivi e l'altezza.

Dimostrare con un controesempio che il numero di connettivi diversi non è una misura di complessità.

Esercizio 2.1.9. Eliminare le parentesi, applicando le convenzioni sulla priorità dei connettivi, dalle seguenti proposizioni:

$$((p) \land ((\neg(q)) \to (\neg(p))))$$
$$((\neg(\neg(\neg(p)))) \lor ((p) \land (q)))$$
$$(((\neg(p)) \lor (\neg(q))) \land ((\neg(p)) \lor (q)))$$
$$(((p) \oplus (\neg(q))) \to ((p) \lor (\neg(q)))).$$

Esercizio 2.1.10. Reintrodurre le parentesi nelle seguenti parole in modo da ottenere, se possibile, proposizioni, o se no spiegare il perché:

$$\neg\neg p$$
$$\neg p \land q \lor r$$
$$p \to q \lor \neg r$$
$$(p \to q) \land p \to q$$
$$p \to q \land p \to q$$
$$p \lor q \land r \to \neg p$$
$$p \land q \land r \lor \neg r$$
$$p \land (\to r \lor q)$$
$$p \oplus \neg q \to \neg p \oplus q$$
$$p \oplus q \lor r.$$

Esercizio 2.1.11. Definire le proposizioni nel seguente modo: Ogni lettera p è una proposizione; se A è una proposizione, anche $\neg(A)$ è una proposizione; se $\bullet$ è un connettivo binario e A e B sono proposizioni, anche $(A) \bullet (B)$ è una proposizione.

Definire il nuovo procedimento per decidere se una parola è una proposizione e costruire l'albero sintattico.

Discutere eventuali vantaggi e svantaggi della definizione alternativa.

2.2 Semantica

La semantica ha a che fare con le interpretazioni, grazie alle quali le proposizioni, con la sostituzione di frasi alle lettere, vengono ad assumere un senso (che a noi non interessa, lo bypassiamo) e diventano vere o false. Tale attribuzione *finale* di valori di verità è per noi l'operazione di interpretazione, che viene studiata in astratto per vedere se abbia proprietà generali, indipendenti dalle interpretazioni concrete.

I valori di verità saranno rappresentati dall'insieme[20] $\{0, 1\}$. Ci si colloca

[20] Altre notazioni per i valori di verità sono $\{\mathbf{False}, \mathbf{True}\}$, $\{F, T\}$, $\{F, V\}$, $\{\bot, \top\}$.

con tale scelta nell'ottica della logica classica a due valori.

Nell'insieme $\{0, 1\}$ è necessario introdurre un minimo di struttura: la più semplice consiste in convenire che $0 < 1$ e usare la sottrazione come se 0 e 1 fossero numeri interi, con $|x|$ a indicare il valore assoluto.

Un'interpretazione è una funzione[21] $i : \mathcal{L} \longrightarrow \{0, 1\}$; una valutazione è una funzione $v : \mathcal{P} \longrightarrow \{0, 1\}$ che soddisfa le seguenti condizioni[22]:

$$\begin{aligned} v((\neg A)) &= 1 - v(A) \\ v((A \wedge B)) &= \min\{v(A), v(B)\} \\ v((A \vee B)) &= \max\{v(A), v(B)\} \\ v((A \oplus B)) &= |v(A) - v(B)| \\ v((A \rightarrow B)) &= \max\{1 - v(A), v(B)\} \\ v((A \leftrightarrow B)) &= 1 - |v(A) - v(B)|. \end{aligned}$$

In alternativa, si considerano 0 e 1 come interi modulo[23] 2, $\{0, 1\} = \mathbb{Z}_2$, e si scrivono le condizioni:

$$\begin{aligned} v((\neg A)) &= 1 + v(A) \\ v((A \wedge B)) &= v(A) \cdot v(B) \\ v((A \vee B)) &= v(A) + v(B) + v(A) \cdot v(B) \\ v((A \oplus B)) &= v(A) + v(B) \\ v((A \rightarrow B)) &= 1 + v(A) \cdot (1 + v(B)) \\ v((A \leftrightarrow B)) &= 1 + (v(A) + v(B)). \end{aligned}$$

Oppure ancora si considera $\{0, 1\}$ come l'algebra di Boole **2**, con le condizioni

$$\begin{aligned} v((\neg A)) &= -v(A) \\ v((A \wedge B)) &= v(A) \circ v(B) \\ v((A \vee B)) &= v(A) + v(B) \\ v((A \oplus B)) &= (v(A) + v(B)) \circ -(v(A) \circ v(b)) \\ v((A \rightarrow B)) &= -v(A) + v(B) \\ v((A \leftrightarrow B)) &= -v(A \oplus B). \end{aligned}$$

---

[21]La notazione con la freccia è usuale in matematica per le funzioni: si intende che a ogni lettera corrisponde un valore di verità, e per la valutazione v che a ogni proposizione corrisponde o 0 o 1.

[22]Si noti che in $v((\neg A))$ e in altre espressioni analoghe ci sono due tipi di parentesi, che andrebbero tipograficamente distinte; quelle interne sono le parentesi della proposizione, quelle esterne servono per la notazione funzionale $v(x)$.

[23]Per chi non sa cosa significa, l'importante è che $1 + 1 = 0$. In pratica i numeri sono divisi in due classi, quella dei pari, rappresentata da 0 e quella dei dispari, rappresentata da 1. La somma di due pari è pari, la somma di due dispari è pari ....

Ogni interpretazione i si estende a una valutazione i^* ponendo

$$i^*((p)) = i(p)$$

e definendo i^* sulle proposizioni composte in modo da soddisfare le condizioni della definizione di valutazione.

Per ogni valutazione v il valore di verità di una proposizione A si ottiene applicando ai valori delle sottoproposizioni immediate di A una funzione, che dipende dal connettivo principale di A.

Tavole di verità

Ad ogni connettivo è associata una *funzione di verità*, cioè una funzione da $\{0,1\}^n$ in $\{0,1\}$, dove n è il numero di argomenti del connettivo ($\{0,1\}^n$ è l'insieme delle n-uple di 0 e 1). Per il loro carattere finito queste funzioni sono rappresentabili mediante tabelle, che sono dette *tavole di verità*.

La tavola di verità della negazione è:

A	$\neg A$
0	1
1	0

la tavola di verità della congiunzione:

A	B	$A \wedge B$
0	0	0
0	1	0
1	0	0
1	1	1

la tavola di verità della disgiunzione:

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

la tavola di verità della disgiunzione esclusiva:

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

la tavola di verità del condizionale:

A	B	$A \to B$
0	0	1
0	1	1
1	0	0
1	1	1

e la tavola di verità del bicondizionale:

A	B	$A \leftrightarrow B$
0	0	1
0	1	0
1	0	0
1	1	1

Quando si deve trovare il valore di verità di una proposizione, o di un numero finito di esse, sotto un'interpretazione, è sufficiente considerare i valori assunti dalle lettere che vi compaiono, quindi le interpretazioni diventano assegnazioni di valori 0 o 1 ad un numero finito di lettere, e per ogni proposizione ce ne è un numero finito. Data una proposizione, il calcolo dei suoi valori di verità per ogni possibile interpretazione si può organizzare in una tabella con i valori progressivi attribuiti alle sottoproposizioni (individuate dall'analisi sintattica), come nei seguenti esempi:

Esempio 2.2.1. Se A è $p \wedge \neg p \to q$:

p	q	$\neg p$	$p \wedge \neg p$	$p \wedge \neg p \to q$
0	0	1	0	1
0	1	1	0	1
1	0	0	0	1
1	1	0	0	1

Esempio 2.2.2. Se A è $p \vee r \to \neg p \wedge (q \to r)$:

p	q	r	$\neg p$	$q \to r$	$\neg p \wedge (q \to r)$	$p \vee r$	A
0	0	0	1	1	1	0	1
0	0	1	1	1	1	1	1
0	1	0	1	0	0	0	1
0	1	1	1	1	1	1	1
1	0	0	0	1	0	1	0
1	0	1	0	1	0	1	0
1	1	0	0	0	0	1	0
1	1	1	0	1	0	1	0

Tali tabelle si chiamano tavole di verità delle proposizioni.

Come si vede dagli esempi, ci sono proposizioni che per ogni interpretazione hanno il valore 1, altre che per alcune interpretazioni hanno il valore 0 e per altre interpretazioni il valore 1. Si possono dare esempi di proposizioni che per ogni interpretazione assumono il valore 0 (esercizio).

Si ricordi che una proposizione, in quanto schema, non è né vera né falsa; solo la sua tavola di verità completa spiega tutti i possibili modi in cui lo schema può realizzarsi nelle diverse interpretazioni.

Esercizi

1. Costruire la tavola di verità delle proposizioni:

$$
\begin{gathered}
(p \to p) \to p \\
p \to (p \to p) \\
p \vee q \to p \wedge q \\
p \vee q \wedge r \to p \wedge r \vee s \\
(p \vee q) \wedge r \to p \wedge (r \vee s) \\
p \to (q \to p).
\end{gathered}
$$

2. Spiegare quale è la disgiunzione usata nella programmazione, in considerazione del fatto che ivi si adotta la valutazione pigra: “quando viene valutata una disgiunzione, e la prima condizione è vera, la seconda condizione non viene esaminata”[24].

3. Trovare le tavole di verità corrispondenti a “a meno che”, “anche se”.

4. Scrivere la tavola di verità per le particelle logiche “né . . . né” e “non (è vero che) sia . . . sia . . . ”.

5. Costruire la tavola di verità per “se . . . allora . . . , altrimenti . . . ”.

   Avvertenza. Si faccia attenzione che il costrutto IF . . . THEN nei linguaggi di programmazione è usato piuttosto come $\leftrightarrow$; se lo statement è falso l’istruzione non viene eseguita: ad esempio “se si esegue

   IF importo $\leq$ saldo THEN saldo := saldo $-$ importo,

   l’enunciato dell’assegnazione verrà eseguito solo se l’importo da prelevare è minore o uguale al saldo”[25].

[24] Horstmann, p. 212.
[25] Horstmann, p. 186.

2.2.1 Validità e conseguenza

Se $i^*(A) = 1$, si dice che A è *vera* nell'interpretazione i, o che i *soddisfa* A, o che i è un *modello* di A, e si scrive anche

$$i \models A.$$

Se esiste almeno una i tale che $i \models A$, si dice che A è *soddisfacibile*, o (semanticamente) *consistente*. Se non esiste alcun modello di A, si dice che A è *insoddisfacibile*, o (semanticamente) *inconsistente*, o *contraddittoria*, o una *contraddizione*. Se per ogni i si ha $i \models A$, si dice che A è *logicamente valida*, o *logicamente vera*, o una *tautologia*, e si scrive

$$\models A.$$

Si dice che B è *conseguenza logica* di A, o che A *implica* B, e si scrive

$$A \models B$$

se per ogni i, se $i \models A$ allora $i \models B$. Si noti che

Osservazione 2.2.3. Per ogni A e B,

$$A \models B \text{ se e solo se } \models A \to B.$$

Siccome $i \models A_1 \wedge \ldots \wedge A_n$ se e solo se $i \models A_j$ per ogni j con $1 \leq j \leq n$, la definizione di modello si può generalizzare dicendo che i soddisfa un insieme di proposizioni T se e solo se $i \models A$ per ogni $A \in T$.

Quindi se A è $A_1 \wedge \ldots \wedge A_n$, invece di $A_1 \wedge \ldots \wedge A_n \models B$ si scrive $\{A_1, \ldots, A_n\} \models B$, o anche $A_1, \ldots, A_n \models B$.

Se $A \models B$ e $B \models A$, si dice che A e B sono *logicamente equivalenti*, o anche solo equivalenti, e si scrive $A \equiv B$.

Osservazione 2.2.4. Per ogni A e B,

$$A \equiv B \text{ se e solo se } \models A \leftrightarrow B.$$

Si noti che $\models$ e $\equiv$ sono segni metalinguistici, non connettivi.

Dalle definizioni semantiche segue immediatamente che

$$A_1, \ldots, A_n \models B \text{ se e solo se } \{A_1, \ldots, A_n, \neg B\} \text{ è insoddisfacibile.}$$

Questo significa che si può assumere come concetto semantico fondamentale sia quello di conseguenza logica sia quello di soddisfacibilità, e a seconda di quale sia privilegiato orientare diversamente la ricerca dei metodi più efficienti per rispondere alle domande semantiche.

La relazione di conseguenza logica è evidentemente transitiva: se $A \models C$ e $C \models B$ allora $A \models B$ (esercizio).

Lo stesso vale per la relazione di equivalenza logica.

Le tautologie, in particolare quelle che sono nella forma di equivalenze o implicazioni, sono dette anche *leggi logiche*.

Un elenco di leggi logiche notevoli è presentato nella tabella 2.1 a pagina 60.

Per verificare queste leggi, dove $A, B, \ldots$ sono qualunque, si devono prima verificare le stesse nel caso particolare che $A, B, \ldots$ siano atomiche (ad esempio $p \to p$ per la legge dell'identità), e poi sfruttare il fatto che se $A[p]$ è una tautologia e B è qualunque, allora anche il risultato della sostituzione di B a p in A è una tautologia (vedi esercizi).

Per le leggi che nella tabella sono scritte come condizionali e non bicondizionali, si vedrà in seguito che l'implicazione inversa in generale non sussiste (salvo alcuni casi, ad esempio per l'inverso della riduzione all'assurdo debole, cioè $\neg A \to (A \to \neg A)$, che rientra nell'affermazione del conseguente).

L'associatività della congiunzione giustifica che si possa scrivere senza ambiguità, indipendentemente dalle convenzioni sulle parentesi, $A \wedge B \wedge C$ per (indifferentemente) $A \wedge (B \wedge C)$ o $(A \wedge B) \wedge C$, o in generale $A_1 \wedge \ldots \wedge A_n$ (e lo stesso per la disgiunzione). $A \wedge (B \wedge C)$ e $(A \wedge B) \wedge C$ sono diverse (si disegni il loro albero sintattico) ma si dice che sono uguali *a meno di* equivalenza logica.

Anche le seguenti sono leggi logiche:

$$\begin{array}{l}
A \to B \leftrightarrow \neg A \vee B \\
(A \leftrightarrow B) \leftrightarrow (A \to B) \wedge (B \to A) \\
A \oplus B \leftrightarrow (A \wedge \neg B) \vee (B \wedge \neg A) \\
A \oplus B \leftrightarrow (A \vee B) \wedge \neg(A \wedge B).
\end{array}$$

Si noti che le due leggi per $\oplus$ forniscono un esempio di come una particella logica possa essere espressa con diversi giri di frase equivalenti; queste equivalenze in genere mostrano cosa significa che frasi diverse vogliono dire la stessa cosa.

Per mezzo di esse, dalle leggi elencate sopra se ne derivano altre; ad esempio dal *modus ponens* e dall'esportazione, con la prima, si ricava

$$A \wedge (\neg A \vee B) \to B \quad \text{sillogismo disgiuntivo.}$$

Ma queste leggi soprattutto permettono di vedere che i connettivi $\oplus, \to, \leftrightarrow$ sono definibili in termini di $\neg, \wedge$ e $\vee$.

$A \to A$	legge dell'identità
$A \leftrightarrow \neg\neg A$	legge della doppia negazione
$A \wedge B \leftrightarrow B \wedge A$	commutatività di $\wedge$
$(A \wedge B) \wedge C \leftrightarrow A \wedge (B \wedge C)$	associatività di $\wedge$
$A \vee B \leftrightarrow B \vee A$	commutatività di $\vee$
$(A \vee B) \vee C \leftrightarrow A \vee (B \vee C)$	associatività di $\vee$
$A \wedge A \leftrightarrow A$	idempotenza di $\wedge$
$A \vee A \leftrightarrow A$	idempotenza di $\vee$
$A \wedge B \to A$	eliminazione di $\wedge$
$A \to A \vee B$	introduzione di $\vee$
$A \wedge (B \vee C) \leftrightarrow (A \wedge B) \vee (A \wedge C)$	distributività
$A \vee (B \wedge C) \leftrightarrow (A \vee B) \wedge (A \vee C)$	distributività
$A \wedge (A \vee B) \leftrightarrow A$	legge di assorbimento
$A \vee (A \wedge B) \leftrightarrow A$	legge di assorbimento
$\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$	legge di De Morgan
$\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$	legge di De Morgan
$\neg A \vee A$	legge del terzo escluso
$\neg(A \wedge \neg A)$	legge di non contraddizione
$A \to B \leftrightarrow \neg B \to \neg A$	legge di contrapposizione
$A \wedge \neg A \to B$	Lewis, o *ex falso quodlibet*
$A \to (B \to A)$	affermazione del conseguente
$\neg A \to (A \to B)$	negazione dell'antecedente
$(A \to B \wedge \neg B) \to \neg A$	legge di riduzione all'assurdo
$(A \to \neg A) \to \neg A$	riduzione all'assurdo debole
$(\neg A \to A) \to A$	consequentia mirabilis
$((A \to B) \to A) \to A$	legge di Peirce
$(A \to B) \vee (B \to A)$	legge di Dummett
$A \to ((A \to B) \to B)$	modus ponens
$A \to (B \to C) \leftrightarrow B \to (A \to C)$	scambio antecedenti
$(A \to C) \wedge (B \to C) \leftrightarrow A \vee B \to C$	distinzione di casi
$(A \to B) \wedge (\neg A \to B) \to B$	distinzione di casi
$(A \to (B \to C)) \to ((A \to B) \to (A \to C))$	distributività di $\to$
$(A \to B) \wedge (B \to C) \to (A \to C)$	transitività di $\to$
$A \to (B \to C) \leftrightarrow (A \wedge B) \to C$	importazione/esportazione delle premesse

Tabella 2.1: Leggi logiche notevoli 1

2.3 Calcolo della deduzione naturale

Alcune leggi sono spesso presentate in forma di regole di *inferenza*; ad esempio il *modus ponens*, invece che da $\models A \to ((A \to B) \to B)$, da

$$\frac{A \quad A \to B}{B};$$

l'eliminazione e l'introduzione della congiunzione, l'introduzione della disgiunzione, il sillogismo disgiuntivo (come eliminazione della disgiunzione) e altre le abbiamo già viste nel capitolo 1.

La regola di introduzione della negazione è giustificata dalla legge di riduzione all'assurdo. Quella di eliminazione della negazione è giustificata dal fatto che $A_1, \ldots, A_n \models B$ equivale all'essere $\{A_1, \ldots, A_n, \neg B\}$ insoddisfacibile. Infatti $\{A_1, \ldots, A_n, \neg B\}$ insoddisfacibile è equivalente a

$$A_1, \ldots, A_n, \neg B \models C \wedge \neg C.$$

Si ritrovano così le regole della deduzione naturale. Le regole e la loro rappresentazione grafica si interpretano semanticamente nel seguente modo. Le leggi corrispondenti permettono di asserire che se sono vere le proposizioni sopra la riga, o *premesse della regola*, allora è vera anche la proposizione sotto la riga, o *conclusione*. Regole d'inferenza di questo genere si dicono corrette se le premesse implicano logicamente la conclusione — quindi le regole sopra elencate sono corrette.

Si chiamano in generale *calcoli logici* i metodi per rispondere ai quesiti logici sulla verità, l'insoddisfacibilità, la conseguenza, metodi che sono procedure guidate dalla sintassi, e che si articolano in applicazioni iterate di regole che producono strutture come sequenze finite o alberi di proposizioni, che si chiamano deduzioni, o *derivazioni*, o *dimostrazioni*.

Il calcolo della deduzione naturale presentato nella sezione 1.2 (quello classico) è un tale calcolo, ideato per provare la sussistenza della validità logica.

Quando si dà un metodo sintattico per rispondere a quesiti di natura semantica (o un calcolo per risolvere un problema), si pone la questione, e la richiesta, della correttezza e completezza del metodo. *Correttezza* significa che le risposte che dà il metodo sono giuste, *completezza* significa che quando la risposta c'è il metodo la dà, quella giusta.

Se le regole del calcolo della deduzione naturale sono, come sono, corrette, è ovvio che:

Teorema 2.3.1 (Correttezza)**.** *Se* $A_1, \ldots, A_n \vdash B$*, allora* $A_1, \ldots, A_n \models B$*.*

Vale anche il viceversa (si veda l'esercizio 2.3.28):

Teorema 2.3.2 (Completezza). *Se* $A_1, \ldots, A_n \models B$, *allora* $A_1, \ldots, A_n \vdash B$.

Per dimostrare che $A_1, \ldots, A_n \models B$ si può allora cercare di dedurre B dalle assunzioni $A_1, \ldots, A_n$.

Si dice peraltro informalmente che una proposizione B si deduce da un'altra A se $A \models B$ e se questo fatto è riconosciuto e certificato da una spiegazione.

Un modo per riconoscere la sussistenza di $A \models B$ è quello di inserire tra A e B altre proposizioni legate tra loro dalla relazione di premesse-conclusione di regole corrette, il che spesso porta a costruire una deduzione.

Ad esempio per stabilire

$$(r \to p \vee q) \wedge r \wedge \neg p \models q$$

si può eseguire il seguente ragionamento: dall'assunzione $(r \to p \vee q) \wedge r \wedge \neg p$ seguono logicamente $r \to p \vee q$, r e $\neg p$ separatamente per eliminazione della congiunzione, quindi $p \vee q$ da $r \to p \vee q$ e r per *modus ponens*, e q da $p \vee q$ e $\neg p$ per il sillogismo disgiuntivo.

Ma questo ragionamento non è altro che la deduzione:

1	$(r \to p \vee q) \wedge r \wedge \neg p$	assunzione
2	$r \to p \vee q$	(E$\wedge$) da 1
3	r	(E$\wedge$) da 1
4	$p \vee q$	(MP) da 2 e 3
5	$\neg p$	(E$\wedge$) da 1
6	q	(E$\vee$) da 4 e 5.

La completezza ci assicura che il calcolo è potente, nel senso che permette di dedurre tutte le tautologie. Non si vorrebbe che fosse troppo potente, cioé che permettesse di dedurre anche altre proposizioni, magari anche alcune contraddittorie, magari tutte. In tal caso il calcolo si direbbe contraddittorio.

Che il calcolo non sia contraddittorio è assicurato dal teorema di correttezza.

La correttezza di un calcolo è utile per poter dimostrare che *non* esiste una derivazione $A_1, \ldots, A_n \vdash B$.

Siccome sappiamo che in generale se $A \vdash B$ allora $A \models B$, allora ad esempio possiamo affermare che $A \vee B \nvdash B$ perché B non è conseguenza logica di $A \vee B$ (e tale proprietà la possiamo verificare in un modo effettivo: $A \vee B \to B$ non è una tautologia).

In un contesto astratto, dare una semantica a un calcolo consiste nell'associare al linguaggio una proprietà delle formule che si conservi dalle premesse alla conclusione nella applicazione di ogni regola. Allora si può dire che se determinate assunzioni hanno tale proprietà, anche ogni formula dedotta da quelle assunzioni deve avere la stessa proprietà (nel caso del calcolo della deduzione naturale per la logica proposizionale e della semantica delle tavole di verità la proprietà è quella di essere soddisfatta in una interpretazione).

2.3.1 Sistemi formali

Facciamo un esempio artificiale, presentando un *sistema formale* in miniatura, cioé un sistema finito di regole per trasformare parole di un alfabeto non interpretato, neanche nelle intenzioni[26].

Le regole sono scritte nella forma $\sigma \rightsquigarrow \tau$, da intendere che la parola σ si può riscrivere come parola τ. Le parole sono rappresentate dalla concatenazione di simboli dell'alfabeto. Usiamo le lettere $x, y, \ldots$ come variabili sulle parole, inclusa la parola vuota.

Alfabeto: **M**, **I**, **U**.

Regole:

$$\begin{array}{lll}
\mathrm{R1}: & x\mathbf{I} & \rightsquigarrow & x\mathbf{IU} \\
\mathrm{R2}: & \mathbf{M}x & \rightsquigarrow & \mathbf{M}xx \\
\mathrm{R3}: & x\mathbf{III}y & \rightsquigarrow & x\mathbf{U}y \\
\mathrm{R4}: & x\mathbf{UU}y & \rightsquigarrow & xy
\end{array}$$

In questo sistema data una parola (come assioma, o assunzione) ci si può chiedere quali parole siano deducibili da essa. Ad esempio da **MI** si derivano le seguenti parole:

1	**MI**	assunzione
2	**MII**	per R2
3	**MIIII**	per R2
4	**MUI**	per R3
5	**MUIU**	per R1
6	**MUIUUIU**	per R2
7	**MUIIU**	per R4
...		

ma anche

1	**MI**	assunzione
2	**MIU**	per R1
...		

[26]L'esempio è preso da D. R. Hofstadter, *Gödel, Escher, Bach*, Basic Books, New York, 1979.

Ci si può chiedere se **MU** sia deducibile da **MI** (o da una qualsiasi altra parola che contenga una sola occorrenza di **I**).

La risposta è negativa in base alla seguente considerazione: per ottenere una parola senza **I** occorre come condizione necessaria partire da una parola che contenga un numero multiplo di 3 di occorrenze di **I** perché l'applicazione delle regole o lascia invariato il numero di occorrenze di **I**, o lo raddoppia o lo riduce di 3.

Sono molto più interessanti questi problemi di impossibilità che non l'esecuzione meccanica, anche se non deterministica, delle regole, che genera le deduzioni.

Esercizi

Esercizio 2.3.3. Verificare con le tavole di verità le leggi logiche elencate in 3.2.

Esercizio 2.3.4. Spiegare in base alla definizione di $\models$ perché

- Se $\models A$ allora $B \models A$ per ogni B.
- Se $\models A$ allora $\models A \vee B$ per ogni B.
- Se $\models A$ e $\models A \to B$ allora $\models B$.

Esercizio 2.3.5. Spiegare perché se $A[p]$ è una tautologia, anche la proposizione che si ottiene sostituendo p con una B qualunque è una tautologia.

Esercizio 2.3.6. Verificare la seguente generalizzazione delle leggi di assorbimento, che $A \equiv A \vee C$ se $C \models A$, e che $A \equiv A \wedge C$ se $A \models C$.

Esercizio 2.3.7. Verificare che $A \equiv T \wedge A$ se T è una tautologia e che $A \equiv F \vee A$ se F è una contraddizione, e dedurlo dal risultato del precedente esercizio.

Esercizio 2.3.8. Verificare che $A \to B \equiv \neg(A \wedge \neg B)$ (sia con le tavole, sia in base alla definizione di interpretazione).

Esercizio 2.3.9. Verificare che $A \oplus B$ è equivalente a $\neg(A \leftrightarrow B)$, in base alla definizione di interpretazione.

Esercizio 2.3.10. Verificare che $\models A \oplus \neg A$.

Esercizio 2.3.11. Trovare un'equivalenza che definisca $\vee$ in termini di $\oplus$ e $\wedge$.

Esercizio 2.3.12. Verificare che $\models A \oplus B \to A \vee B$ ma non viceversa.

Esercizio 2.3.13. Spiegare perché $A \to A \oplus B$ non è logicamente vera.

Esercizio 2.3.14. Verificare che $p \vee q$ è equivalente a $p \vee (q \wedge \neg p)$ ed a $p \oplus (q \wedge \neg p)$.

Esercizio 2.3.15. Verificare che la regola del sillogismo disgiuntivo è corretta anche con $\oplus$ al posto di $\vee$.

Esercizio 2.3.16. Notare che $\neg(A \oplus B) \equiv \neg A \oplus B$ (provare a trovare frasi in italiano che si possono dire bene in entrambi i modi, e dove sia chiaro che la disgiunzione è forte — non è facile per $\neg(A \oplus B)$).

Esercizio 2.3.17. Verificare se valgono le distributive tra $\wedge$ e $\oplus$.

Esercizio 2.3.18. Verificare se valgono analoghe di De Morgan per $\oplus$ e $\wedge$.

Esercizio 2.3.19. Verificare se $A \oplus (B \oplus C) \equiv (A \oplus B) \oplus C$.

Esercizio 2.3.20. In base al precedente esercizio, discutere quando $A_1 \oplus \ldots \oplus A_n$ è vera.

Esercizio 2.3.21. Verificare che $\neg\neg\neg\neg A \equiv A$ e che $\neg\neg\neg\neg\neg\neg\neg A \equiv \neg A$. Generalizzare.

Esercizio 2.3.22. Si consideri il problema del *merging* di due liste List1 e List2 in una terza lista List3 (ad esempio nomi, in ordine alfabetico).

Una prima formulazione dell'algoritmo è la seguente: nello scorrere le due liste, se List1 non è esaurita e List2 è esaurita oppure l'elemento in considerazione di List1 precede il primo non ancora inserito di List2, allora l'elemento di List1 è inserito in List3.

Un'altra formulazione potrebbe essere la seguente: il prossimo elemento in List3 è preso da List1, ed è il primo elemento di List1 non ancora inserito, quando List1 non è esaurita e List2 sì, oppure quando List1 non è esaurita e l'elemento in considerazione di List1 precede il primo non ancora inserito di List2.

Usando lettere p, q, r per rappresentare rispettivamente "List1 non è esaurita", "List2 è esaurita" e "l'elemento di List1 precede quello di List2", scrivere le proposizioni corrispondenti alle due versioni delle condizioni (che portano entrambe a mettere in List3 l'elemento in esame di List1), e discutere se siano o no equivalenti, in base a quali leggi.

Si noti che ovviamente nella prima versione c'è un'ambiguità dovuta alla presenza di congiunzione e disgiunzione; discutere le due versioni e scegliere quella giusta.

Esercizio 2.3.23. Si distribuiscono carte da gioco, più di una per giocatore, e si sa che un giocatore ha in mano un Asso o un Re. Si considerino le seguenti due proposizioni:

A: se il giocatore in mano ha un Asso, ha un 2
B: se il giocatore in mano ha un Re, ha un 2.

Che cosa si può dedurre da $A \oplus B$, cioè se esattamente una tra le proposizioni A e B è vera?

Che cosa si può dedurre se entrambe le proposizioni A e B sono vere?

Esercizio 2.3.24. Per conquistare la principessa, Aladino deve scegliere di aprire una di due scatole A e B; sa che nelle scatole ci sono o un anello (segno di fidanzamento) o un serpente velenoso che uccide all'istante chi apre le scatole. Potrebbero anche esserci due serpenti, se il Vizir non lo vuole come genero, o due anelli, se al contrario lo apprezza. Sulla scatola A è scritto: "Almeno una di queste scatole contiene un anello"; sulla scatola B è scritto: "Nella scatola A c'è un serpente velenoso". Ad Aladino viene detto che o entrambe le scritte sono vere, o entrambe false. Quale scatola apre?

Esercizio 2.3.25. "Se io ho ragione, tu hai torto; se tu hai ragione, io ho torto; quindi uno di noi ha ragione". Corretto o no? Perché?

Esercizio 2.3.26. "La storia insegna che non si impara niente dalla storia". Vero o falso? Perché?

Suggerimento. Riduzione all'assurdo debole.

Esercizio 2.3.27. Trovare una deduzione naturale per tutte le leggi logiche notevoli della tabella 2.1 a pagina 60.

Esercizio 2.3.28. Dimostrare la completezza della deduzione naturale nella forma:

$$\text{se} \models A \text{ allora } \vdash A.$$

Suggerimento. Per ogni interpretazione i fissata, si indichi con p^i la proposizione p o rispettivamente $\neg p$ a seconda che $i(p) = 1$ o $i(p) = 0$.

Si dimostri, per induzione sull'altezza di $A[p_1, \ldots, p_n]$, che per ogni i

$$\text{se } i \models A \text{ allora } p_1^i, \ldots, p_n^i \vdash A$$

e

$$\text{se } i \not\models A \text{ allora } p_1^i, \ldots, p_n^i \vdash \neg A.$$

Si considerino quindi due interpretazioni i_1 e i_2 tali che $i_1(p_j) = i_2(p_j)$ per $j = 1, \ldots, n-1$, e $i_1(p_n) \neq i_2(p_n)$. Allora, indicando con i le due interpretazioni coincidenti, o meglio le restrizioni di i_1 e i_2 a $\{p_1, \ldots, p_{n-1}\}$,

$$p_1^i, \ldots, p_{n-1}^i, p_n \vdash A$$

e

$$p_1^i, \ldots, p_{n-1}^i, \neg p_n \vdash A,$$

perché A vale sempre 1, quindi

$$p_1^i, \ldots, p_{n-1}^i \vdash A.$$

Si iteri, scegliendo una nuova interpretazione che differisce da i solo su p_{n-1}, e così via finché non si ottiene $\vdash A$.

2.3.2 Appendice: Sull'implicazione

Abbiamo distinto il condizionale, che è un connettivo, o il nome di una proposizione della forma $A \to B$, dall'implicazione, che è una relazione tra proposizioni, e non si scrive $A \to B$ ma $\models A \to B$. "A implica B" significa "il condizionale $A \to B$ è una tautologia".

La terminologia è qualche volta ambigua perché per leggere ad esempio una regola come il sillogismo disgiuntivo si trova anche detto "se A e $\neg A \vee B$ allora B", in alternativa a "A e $\neg A \vee B$ implicano B". Se si è in un contesto deduttivo si capisce forse che si sta parlando dell'implicazione e non leggendo semplicemente la forma di una proposizione. L'importante ad ogni modo non è la terminologia quanto capire la differenza.

Il soggetto di "A implica B" non è A ma $A \to B$. Qualche volta, in analogia al caso dell'equivalenza, si introduce un simbolo speciale per l'implicazione, che assomigli a un connettivo, ad esempio $A \Rightarrow B$; il nostro simbolo è $\models$.

Si dice ad esempio "il condizionale $p \to p \vee q$ ha cinque simboli", non "l'implicazione $p \to p \vee q$ ha cinque simboli", perché l'implicazione è un fatto che sussiste o no, e un fatto non è Êformato da simboli. Al massimo è un predicato, sotto cui cadono alcuni condizionali, come in "il condizionale $p \to p \vee q$ è un'implicazione". Oppure si può dire che vale l'implicazione $p \to p \vee q$, ma non si parlerà ad esempio dell'implicazione $p \to q \vee r$, che non è una tautologia.

Siccome purtroppo la terminologia non è uniforme, e si possono trovare usate entrambe le parole, bisogna fare attenzione al contesto.

Nella tradizione logica, il condizionale era anche chiamato “implicazione materiale”, per distinguere la relazione di conseguenza da altre forme di implicazione, o da altri sensi del costrutto “se ... allora”.

In effetti, il significato di “se ... allora” è polimorfo:

- significato logico (o inferenziale):

  Se tutti gli uomini sono mortali e Socrate è un uomo, allora Socrate è mortale.

- significato definitorio:

  Se è scapolo, allora non è sposato.

- significato causale:

  Se si immerge una cartina di tornasole e diventa rossa, allora il liquido è un acido.

- significato materiale:

  Se la Terra vola, allora la Terra è piatta.

È difficile trovare qualcosa di positivo in comune tra queste diverse accezioni del “se ... allora”. In particolare il caso che ha sollevato maggiori discussioni è l'ultimo, come considerare il condizionale se antecedente e conseguente sono entrambe false.

Una cosa in comune ce l'hanno, ed è che in tutte le accezioni l'unico modo per dichiarare il condizionale falso è quello di riscontrare antecedente vera e conseguente falsa, anche per il significato materiale: “se la Terra è rotonda, allora il Sole è freddo” si considera falso.

Allora il significato parziale comune si può esprimere riempiendo la tavola di verità con i valori che sono di fatto quelli di $\neg(A \wedge \neg B)$:

> Un condizionale è corretto [secondo Crisippo] se la negazione della sua conclusione è incompatibile con la sua premessa (Sesto Empirico, *Schizzi pirroniani*, II, 110-2).

Si ottiene così quella che gli antichi chiamavano implicazione materiale:

> Secondo lui [Filone di Megara] ci sono tre modi in cui un condizionale può essere vero, e uno in cui può essere falso. Perché un condizionale è vero quando inizia con una verità e termina con una verità, come “se è giorno, è chiaro”. Ed è vero anche quando inizia con una falsità e termina con una falsità, come “se la terra vola, la terra ha ali”. Analogamente, è vero un condizionale che

> inizia con una falsità e termina con una verità, come "se la terra vola, la terra esiste". Un condizionale è falso soltanto quando inizia con una verità e termina con una falsità, come "se è giorno, è notte" (Sesto Empirico, *Contro i matematici*, VIII, 113).

Con questa scelta per la tavola di $\rightarrow$ si giustifica la regola del *modus ponens*, che è quello che interessa, per l'uso che se ne fa nei discorsi con "se ... allora".

Il motivo per cui il condizionale è difficile e controverso è che non gli si può associare una rappresentazione mentale immediata di quello che descrive. Quando si ascolta $A \wedge B$, le rappresentazioni nella mente del fatto descritto da A e di quello descritto da B vengono fuse in un'unica rappresentazione, del fatto descritto da $A \wedge B$, affiancandole o integrandole; anche con $A \vee B$ le due rappresentazioni possono essere compresenti, con l'attenzione che si sposta dall'una all'altra e viceversa, come se si guardassero alternativamente due quadri vicini. Con il condizionale non è possibile avere una rappresentazione del fatto descritto da $A \rightarrow B$, combinando quelle relative ad A e B. Non esiste una rappresentazione unica della falsità di A. Vengono meno perciò gli ausili dell'immaginazione e della sensibilità; l'unico modo per dominare il condizionale è quello di imparare bene fino a interiorizzarle le sue condizioni d'uso, sia il calcolo dei valori di verità sia le leggi e le regole che lo concernono.

La definizione del condizionale tuttavia non è solo adeguata per svolgere le dimostrazioni, grazie alla giustificazione del *modus ponens*, ma è anche comoda (nella scelta di dare il valore vero quando l'antecedente è falsa) per la costruzione generale dei linguaggi formali, e la trattazione dei quantificatori universali ristretti, come abbiamo visto.

Capitolo 3

Insiemi e algebre di Boole

3.1 Insiemi

Esempi di frasi a cui si applicano utilmente le nozioni e le tecniche logiche sono le formule matematiche; gli studenti sono abituati alle formule senza quantificatori, anzi sono queste le uniche che (probabilmente hanno visto e che) sono abituati a chiamare formule. In esse compaiono le variabili x, y,...

Sono le formule che abbiamo detto sono da evitare, per la loro ambiguità, quando si deve formalizzare un discorso.

Di una formula come $1 < x < 3$ non si può dire che è né vera né falsa, in quanto non è precisato se la variabile debba essere intesa in senso universale o particolare. Meglio, non la si può pensare come una frase che trasmette un'informazione. Non la si riesce quasi neanche a leggere in italiano: "un numero compreso tra 1 e 3"? "i numeri compresi tra 1 e 3"? ma manca il verbo.

Di solito la si legge come "x è compreso tra 1 e 3", cioè come una proposizione che afferma qualcosa a proposito di x, dove x denota un elemento non precisato dell'universo; questa lettura, con l'idea di un nome per un elemento non precisato dell'universo, è fonte di equivoci difficilmente rimediabili sulla natura delle variabili, una volta che si sia imposta.

Infatti, come leggere allora $3 < x < 1$? Di fronte a questa si tende a dire che non la si dovrebbe scrivere, perché "non esiste un numero maggiore di 3 e minore di 1", ma allora si fa una affermazione sulla formula, si afferma che dovrebbe essere scritta come $\neg\exists x(3 < x < 1)$. Non si può pretendere infatti che $3 < x < 1$ *stia per* $\neg\exists x(3 < x < 1)$, è solo una scrittura sbagliata.

Per interpretare $1 < x < 3$ come una frase che trasmette un'informazione, dobbiamo immaginare questo teatro: è come se fosse presente un numero incappucciato, che dice “io sono compreso tra 1 e 3”. Se si toglie il cappuccio ed appare 2 ha detto il vero, se appare 0, o 3 o 5 ha detto il falso.

Supponiamo che l’universo sia costituito dai numeri naturali. Se il numero incappucciato continua dicendo “quindi io sono il numero 2”, bisogna ammettere che l’inferenza è corretta, anche senza togliergli il cappuccio. La formula $1 < x < 3$ è soddisfatta dal solo elemento 2, e possiamo affermare $1 < x < 3 \to x = 2$, intendendo che $\forall x(1 < x < 3 \to x = 2)$ è vero.

Se invece l’universo di discorso, che dalla formula in sé non si evince, è quello dei numeri reali, la formula è soddisfatta anche da 1,1, da 1,9, da 2 ,5, da $\sqrt{2}$ e da tutti gli infiniti elementi dell’intervallo $(1, 3)$. In questo caso $\forall x(1 < x < 3 \to x = 2)$ non è vero.

Le frasi con variabili senza quantificatori, quando sono usate correttamente, ma in generale all’interno di discorsi più ampi, intendono le variabili in senso universale. Le variabili usate nella esperienza scolastica con l’algebra sono usate in questo modo.

Tuttavia la matematica è diversa dal linguaggio comune, e in essa tali formule una svolgono un’altra utile funzione, che non è se non raramente presente nel discorso naturale, quella di definire vari insiemi di enti.

La lettura che invece ci mette sulla strada per capire la funzione autonoma di una formula come $1 < x < 3$ è la seguente: “i numeri compresi tra 1 e 3”. La formula non asserisce alcunché di fattuale, essa *definisce* l’insieme dei numeri compresi tra 1 e 3, insiemi che sono diversi nel contesto rispettivamente dei numeri naturali, razionali o reali. Fissata tale ottica, potremmo dire che la formula denota l’intervallo (1,3).

Le formule $A[x]$ intervengono dunque nello studio degli insiemi definibili; oltre a definire insiemi, possono servire anche a fare affermazioni su di essi. Ad esempio per affermare che l’insieme definito da $A[x]$ è contenuto nell’insieme definito da $B[x]$ si può scrive $\forall x(A[x] \to B[x])$, o qualcosa di più preciso che indichi anche l’universo di discorso.

Consideriamo ora un argomento nel quale possiamo da una parte applicare quanto imparato a proposito dei connettivi, e dall’altra iniziare a familiarizzarci con le variabili: le variabili sono usate o per fare affermazioni universali, anche senza la scrittura esplicita di $\forall$, o per definire insiemi.

Se X è l’insieme degli elementi di un universo che soddisfano la proprietà definita da $A[x]$, vale

$$\forall x(x \in X \leftrightarrow A[x]),$$

dove $x \in X$ si legge “x appartiene a X”, o “x è un elemento di X”; $x \notin X$ significa che x non appartiene a X, è un’abbreviazione per $\neg(x \in X)$.

Il simbolo X viene introdotto in corrispondenza alla formula $A[x]$.

Dal punto di vista grammaticale l'affermazione $x \in X$ è la attribuzione a x di una proprietà X, e la notazione logica sarebbe $X(x)$, ma la scrittura $x \in X$ è standard.

Gli insiemi si possono identificare con le proprietà nel senso che a ogni proprietà corrisponde l'insieme degli individui che godono di quella proprietà, e a ogni insieme corrisponde la proprietà di appartenere a quell'insieme.

Se Y è l'insieme definito da $B(x)$, la formula $\forall x(A[x] \rightarrow B[x])$ diventa $\forall x(x \in X \rightarrow z \in Y)$.

Se ci si limita ad affermazioni come questa di tipo universale, si possono tralasciare i quantificatori e scrivere $x \in X \rightarrow x \in Y$, ottenendo formule composte con i connettivi le quali risultano vere o false ogni volta che a x si assegna un particolare individuo dell'universo di discorso.

3.1.1 Algebra degli insiemi

Consideriamo un dominio di discorso costituito da un insieme U, "U" per "universo", e un linguaggio che abbia senso per U, cioè i cui simboli di funzioni o relazione abbiano un corrispondente in funzioni e relazioni su U. Supponiamo[1] che ogni elemento a di U abbia un nome c_a nel linguaggio.

Allora a ogni formula $A[x]$ è associato un insieme, l'insieme definito da $A[x]$ in U, che si può chiamare *insieme di verità* di $A[x]$ in U:

$$V_{A[x]} = \{a \in U \mid A[c_a] \text{ è vera in } U\},$$

che si legge "l'insieme degli elementi a di U per cui $A[c_a]$ è vera in U", o anche l'insieme degli elementi di U che soddisfano la condizione $A[x]$ in U.

Talvolta si scrive anche $\{x \in U \mid A[x]\}$ o addirittura $\{x \mid A[x]\}$ se è chiaro anche l'insieme ambiente U, ma la notazione è ambigua e ingannevole, perché x viene a indicare sia un simbolo sia un elemento di U.

Una volta chiarito il concetto, si può continuare a usare la notazione abituale, essendo consapevoli tuttavia che non ha senso sostituire un oggetto, per esempio un numero reale, al posto di una variabile in una formula.

Se $A[x]$ contiene quantificatori, per vedere se è vera in U bisogna restringere a U i quantificatori stessi. Ad esempio $\{x \in U \mid \exists y(y^2 = x) \text{ in } U\}$ è lo stesso di $\{x \in U \mid \exists y(y \in U \wedge (y^2 = x))\}$ (e questa formula se U è l'insieme dei numeri reali definisce l'insieme dei numeri positivi, mentre se U è l'insieme dei numeri razionali definisce l'insieme dei quadrati, che tuttavia non comprende tutti i numeri positivi, ad esempio non 2).

[1] Si tratta di una soluzione tecnica, non l'unica possibile, ma inizialmente la più comoda, per dare un senso rigoroso a "assegnare a x un particolare individuo dell'universo di discorso".

Se $A[x]$ è una A che non contiene x, ad esempio $\exists y(y^2 = 2)$, allora $A[x]$ è vera o falsa in U indipendentemente dalla scelta degli elementi considerati, per cui se è vera, in $\{x \in U \mid A\}$ ci sono tutti gli elementi di U; se è falsa, nessuno.

Con la notazione $\{x_1, \ldots, x_n\}$ si indica l'insieme i cui elementi sono x_1, $\ldots$, x_n.

L'insieme $\{x, y\}$ si chiama coppia (non ordinata) di x e y, che sono gli unici elementi di $\{x, y\}$: $x \in \{x, y\}$ e $y \in \{x, y\}$[2], e inoltre

$$z \in \{x, y\} \to z = x \vee z = y.$$

Con questa scrittura intendiamo che $\forall x \forall y \forall z(z \in \{x, y\} \to z = x \vee z = y)$ è vera in U, e nel seguito useremo tale forma di abbreviazione.

La coppia $\{x, y\}$ ha due elementi se $x \neq y$; altrimenti se $x = y$ ne ha uno solo, si indica $\{x\}$ e si chiama anche insieme unitario, o *singoletto* di x.

Esempio 3.1.1. Se $U = \{a, b, c, d\}$ e $A[x]$ è $x = a \vee x = b$, l'insieme definito da $A[x]$ in U è $\{a, b\}$.

Se U è l'insieme dei numeri naturali e $A[x]$ è la condizione "x è divisibile per 2", l'insieme di verità di $A[x]$ è l'insieme dei numeri pari, e tale insieme è definito dalla condizione "x è divisibile per 2".

Un insieme di verità è un *sottoinsieme* di U; si dice che X è un sottoinsieme di Y, o che è contenuto[3] in Y, in simboli $X \subseteq Y$, se ogni elemento di X è anche elemento di Y: per ogni x, se $x \in X$ allora $x \in Y$.

D'ora in avanti usiamo le lettere $X, Y, \ldots$ per indicare sottoinsiemi di un insieme U, indipendentemente dalle formule in corrispondenza alle quali sono introdotte per denotare gli insiemi di verità di quelle formule.

Qualche volta, raramente, si scrive $Y \supseteq X$ per $X \subseteq Y$.

Si dice che X è un sottoinsieme proprio di Y, e si scrive $X \subset Y$, oppure $X \subsetneq Y$, se $X \subseteq Y$ ma $X \neq Y$.

Se $X \subseteq Y$ e $Y \subseteq X$ allora X e Y hanno gli stessi elementi; questo per definizione significa che $X = Y$. Quello che caratterizza gli insiemi non sono le loro eventuali definizioni ma i loro elementi; ad esempio l'insieme dei triangoli con tre lati uguali e l'insieme dei triangoli con tre angoli uguali sono lo stesso insieme. Così pure $\{x, y\} = \{y, x\}$, da cui la dizione "non ordinata" per la coppia.

Le operazioni insiemistiche principali, sui sottoinsiemi di un insieme U, sono le seguenti:

[2]Talvolta si scrive $x, y \in X$ per "$x \in X$ e $y \in X$", quindi $x, y \in \{x, y\}$.

[3]Si distingue tra "essere contenuto in un insieme", che si riferisce a sottoinsiemi, ed "appartenere a un insieme", che si riferisce ad elementi.

Complemento Il complemento di X (rispetto a U) è l'insieme degli elementi di U che non appartengono a X:

$$\sim X = \{x \in U \mid x \notin X\}.$$

Differenza La differenza di X meno Y è l'insieme degli elementi di U che appartengono a X e non a Y:

$$X \setminus Y = \{x \in U \mid x \in X \wedge x \notin Y\}.$$

Differenza simmetrica La differenza simmetrica di X e Y è l'insieme degli elementi di U che appartengono a X e non a Y o a Y e non a X:

$$X \triangle Y = \{x \in U \mid x \in X \oplus x \in Y\}.$$

Intersezione L'intersezione di X e Y è l'insieme degli elementi di U che appartengono sia a X sia a Y:

$$X \cap Y = \{x \in U \mid x \in X \wedge x \in Y\}.$$

$X \cap Y$ si legge: "X intersezione Y o "X intersecato con Y o "l'intersezione di X e Y.

Unione L'unione di X e Y è l'insieme degli elementi di U che appartengono ad almeno uno dei due insiemi X e Y:

$$X \cup Y = \{x \in U \mid x \in X \vee x \in Y\}$$

$X \cup Y$ si legge: "X unione Y o "X unito a Y o "l'unione di X e Y.

L'intersezione di X e Y è il più grande insieme che è contenuto sia in X sia in Y, nel senso che[4]

$$X \cap Y \subseteq X \quad \text{e} \quad X \cap Y \subseteq Y$$

e

$$\text{se } Z \subseteq X \text{ e } Z \subseteq Y \text{ allora } Z \subseteq X \cap Y$$

mentre l'unione di X e Y è il più piccolo insieme che contiene sia X sia Y, nel senso che

$$X \subseteq X \cup Y \quad \text{e} \quad Y \subseteq X \cup Y$$

[4]Si noti che non occorrono parentesi perché non è possibile interpretare questa formula come $X \cap (Y \subseteq X)$ in quanto si avrebbe un'operazione tra un insieme e una asserzione — un errore di tipo, si dice in logica. Qualche volta le parentesi si mettono per agevolare la lettura.

e

$$\text{se } Y \subseteq X \text{ e } Z \subseteq X \text{ allora } Y \cup Z \subseteq X.$$

Per dimostrare $X \cap Y \subseteq X$ ad esempio, bisogna dimostrare che

$$x \in X \cap Y \to x \in X$$

è sempre vera in U.

Se si elimina $x \in X \cap Y$ sostituendola con $x \in X \wedge x \in Y$, equivalente per la definizione di $\cap$, si ha

$$x \in X \wedge x \in Y \to x \in X,$$

dove occorrono solo connettivi e formule del tipo $x \in X, x \in Y$. Sostituendo queste frasi, non ulteriormente analizzabili, con lettere proposizionali, si ottiene

$$p \wedge q \to p$$

che è una tautologia (e possiamo chiamare tautologia anche $x \in X \wedge x \in Y \to x \in X$). Poiché per ogni $a \in U$ le formule $c_a \in X$, $c_a \in Y$ diventano vere o false, ogni scelta di un a si può pensare come una interpretazione delle lettere proposizionali. Siccome la proposizione è una tautologia, la frase è vera per ogni $a \in U$.

Oppure, dal punto di vista deduttivo, possiamo osservare che $\vdash x \in X \wedge x \in Y \to x \in X$. In seguito useremo sempre implicitamente questo ragionamento[5]

Per dimostrare la massimalità dell'intersezione, cioè che se $Z \subseteq X$ e $Z \subseteq Y$ allora $Z \subseteq X \cap Y$, si deve far vedere che da

$$x \in Z \to x \in X$$

e da

$$x \in Z \to x \in Y$$

segue

$$x \in Z \to x \in X \cap Y$$

ovvero

$$x \in Z \to x \in X \wedge x \in Y.$$

Ma

$$p \to q, p \to r \models p \to q \wedge r.$$

[5]E in casi analoghi scriveremo semplicemente $x \in X \wedge x \in Y \to x \in X$ omettendo il segno di validità logica $\models$ o quello di derivalibilità $\vdash$.

Si faccia attenzione che in italiano l'unione è descritta anche con la congiunzione: l'unione di X e di Y contiene gli elementi di X *e* di Y.

Nella proprietà di minimalità dell'unione troviamo la spiegazione dello scambio di "e" ed "o" osservato in precedenza in certe frasi. Se si indica con Y l'insieme delle mele, con Z l'insieme delle pere, e con X l'insieme dei frutti, allora la frase "mele e pere sono frutti", intesa come "le mele sono frutti e le pere sono frutti" significa che $Y \subseteq X \wedge Z \subseteq X$, ma questa implica $Y \cup Z \subseteq X$, cioè che "mele o pere sono frutti".

Viceversa, se $Y \cup Z \subseteq X$, allora siccome $Y \subseteq Y \cup Z$ si ha, per la transitività di $\subseteq$ — vedi oltre — che $Y \subseteq X$ e analogamente $Z \subseteq X$, cioè "mele o pere sono frutti" implica a sua volta "le mele sono frutti e le pere sono frutti".

Le operazioni insiemistiche corrispondono ai connettivi: l'appartenenza al complemento è definita mediante la negazione, l'appartenenza all'intersezione mediante la congiunzione, e così via.

Viceversa, ai connettivi proposizionali corrispondono le operazioni insiemistiche sugli insiemi di verità delle proposizioni componenti.

$$\begin{aligned} V_{\neg A[x]} &= \sim V_{A[x]} \\ V_{A[x] \wedge B[x]} &= V_{A[x]} \cap V_{B[x]} \\ V_{A[x] \vee B[x]} &= V_{A[x]} \cup V_{B[x]}. \end{aligned}$$

In particolare si ha $V_{x \in X} = X$.

Si può osservare allora che le operazioni non sono tutte indipendenti, ad esempio:

$$X \setminus Y = X \cap (\sim Y).$$

Infatti

$$\begin{aligned} X \setminus Y &= \{x \mid x \in X \wedge x \notin Y\} \\ &= \{x \mid x \in X \wedge x \in \sim Y\} \\ &= \{x \mid x \in X \cap (\sim Y)\} \\ &= X \cap (\sim Y). \end{aligned}$$

Ma le mutue relazioni delle operazioni le vedremo meglio più avanti.

Per semplificare la scrittura, ed evitare alcune coppie di parentesi, si può convenire su un ordine di priorità dei simboli di operazione $(\sim, \cap, \cup)$, analogo a quello dei connettivi. Così $X \cap (\sim Y)$ si potrebbe scrivere $X \cap \sim Y$. Tuttavia anche se la convenzione è sempre valida, qualche volta scriveremo parentesi che sarebbero inutili, per comodità di lettura.

L'insieme vuoto $\emptyset$ è l'insieme che non ha alcun elemento, ed è un sottoinsieme di qualsiasi U, definito da una condizione contraddittoria qualunque:

$$\emptyset = \{x \in U \mid A[x] \wedge \neg A[x]\},$$

o

$$\emptyset = \{x \in U \mid x \neq x\}.$$

Se si denotasse questo insieme $\emptyset_U$ e si definisse $\emptyset_V = \{x \in V \mid x \neq x\}$ si avrebbe $\emptyset_U = \emptyset_V$ perché i due insiemi hanno gli stessi elementi, nessuno per entrambi.

Caratteristica dell'insieme vuoto è che per ogni x, in qualunque U, $x \notin \emptyset$.

Due insiemi X e Y la cui intersezione sia vuota, $X \cap Y = \emptyset$, cioè non abbiano alcun elemento in comune, si dicono *disgiunti*.

Le relazioni tra le operazioni insiemistiche sono espresse da diverse leggi, un certo numero delle quali è elencato qui di seguito.

1	$X \cap X = X$	idempotenza dell'intersezione
2	$X \cup X = X$	idempotenza dell'unione
3	$X \cap Y = Y \cap X$	commutatività dell'intersezione
4	$X \cup Y = Y \cup X$	commutatività dell'unione
5	$X \cap (Y \cap Z) = (X \cap Y) \cap Z$	associatività dell'intersezione
6	$X \cup (Y \cup Z) = (X \cup Y) \cup Z$	associatività dell'unione
7	$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$	distributività di $\cap$ rispetto a $\cup$
8	$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$	distributività di $\cup$ rispetto a $\cap$
9	$X \cap (X \cup Y) = X$	assorbimento
10	$X \cup (X \cap Y) = X$	assorbimento
11	$\sim(\sim X) = X$	doppio complemento
12	$\sim(X \cap Y) = (\sim X) \cup (\sim Y)$	legge di De Morgan
13	$\sim(X \cup Y) = (\sim X) \cap (\sim Y)$	legge di De Morgan
14	$\sim \emptyset = U$	
15	$\sim U = \emptyset$	
16	$X \cap (\sim X) = \emptyset$	legge dell'inverso per$\cap$
17	$X \cup (\sim X) = U$	legge dell'inverso per $\cup$
18	$X \cap U = X$	legge dell'elemento neutro per $\cap$
19	$X \cup U = U$	
20	$X \cap \emptyset = \emptyset$	
21	$X \cup \emptyset = X$	legge dell'elemento neutro per $\cup$.

Esistono altre leggi che riguardano la relazione $\subseteq$ (alcune già menzionate), come

22	$X \subseteq X$
23	$\emptyset \subseteq X$
24	$X \subseteq U$
25	$X \subseteq X \cup Y$
26	$X \cap Y \subseteq X$

e proprietà come

27 se $X \subseteq Y$ e $Y \subseteq Z$ allora $X \subseteq Z$
28 $X \subseteq Y$ se e solo se $X \cap Y = X$
29 $X \subseteq Y$ se e solo se $X \cup Y = Y$
30 $X \subseteq Y$ se e solo se $X \cap (\sim Y) = \emptyset$
31 $X \subseteq Y$ se e solo se $\sim X \cup Y = U$
32 se $X \subseteq Y$ e $X \subseteq Z$ allora $X \subseteq (Y \cap Z)$
33 se $Y \subseteq X$ e $Z \subseteq X$ allora $(Y \cup Z) \subseteq X$.

Ma non tutte sono indipendenti. La loro dimostrazione può consistere nel mostrare direttamente che i due insiemi implicati hanno gli stessi elementi.

Esempi

3 $X \cap Y = Y \cap X$.

Dimostrazione. Se $x \in X \cap Y$, allora $x \in X \wedge x \in Y$; ma per la commutatività della congiunzione si ha allora $x \in Y \wedge x \in X$, quindi $x \in Y \cap X$. Il viceversa, partendo da $x \in Y \cap X$, è analogo. □

4 $X \cup Y = Y \cup X$.

Dimostrazione. Se $x \in X \cup Y$ allora $x \in X \vee x \in Y$. La conclusione segue come sopra per la commutatività della disgiunzione. Oppure usiamo la distinzione di casi: se $x \in X$, allora $x \in Y \vee x \in X$ per introduzione della disgiunzione. Se $x \in Y$ allora pure $x \in Y \vee x \in X$. Quindi $x \in X \vee x \in Y \to x \in Y \vee x \in X$ e $X \cup Y \subseteq Y \cup X$. Il viceversa è analogo. □

5 $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$.

Dimostrazione. Mostriamo prima che $X \cap (Y \cup Z) \subseteq (X \cap Y) \cup (X \cap Z)$. Se $x \in X \cap (Y \cup Z)$ allora $x \in X$ e $x \in Y \cup Z$. Ci sono due casi: o $x \in Y$ o $x \in Z$. Nel primo caso, $x \in X$ e $x \in Y$, quindi $x \in X \cap Y$, e quindi $x \in (X \cap Y) \cup (X \cap Z)$ per la 25, che supponiamo dimostrata[6]. Nel secondo caso, $x \in X$ e $x \in Z$, quindi $x \in X \cap Z$ e quindi x appartiene a $(X \cap Y) \cup (X \cap Z)$, per la 25 e la 4.

Si mostri ora nello stesso modo (esercizio) che $(X \cap Y) \cup (X \cap Z) \subseteq X \cap (Y \cup Z)$, e l'uguaglianza è provata. □

[6] La dimostrazione è implicita nella precedente dimostrazione di 4.

21 $X \cup \emptyset = X$.

Dimostrazione. Se $x \in X \cup \emptyset$, allora $x \in X \vee x \in \emptyset$, ma $x \notin \emptyset$ quindi per eliminazione della disgiunzione $x \in X$. Il viceversa segue dalla 25. □

24 $X \subseteq U$.

Dimostrazione. $x \in U \to (x \in X \to x \in U)$ (esercizio). □

23 $\emptyset \subseteq X$.

Dimostrazione. Per ogni x, $x \in \emptyset \to x \in X$ segue da $x \notin \emptyset$, qualunque sia X. □

17 $X \cup (\sim X) = U$.

Dimostrazione. Per ogni x, $x \in X \vee \neg(x \in X)$. Così si dimostra $\supseteq$, il viceversa è 24. □

30 $X \subseteq Y$ se e solo se $X \cap (\sim Y) = \emptyset$.

Dimostrazione. Da sinistra a destra. Se $x \in X$ allora $x \in Y$; se ora esistesse un $x \in X \cap (\sim Y)$ si avrebbe una contraddizione $x \in Y$ e $x \in \sim Y$. La dimostrazione è per assurdo: per dimostrare $A \to B$ si assume A in vista della introduzione del condizionale, quindi si assume $\neg B$ in vista dell'eliminazione della negazione. Si ottiene una contraddizione, quindi B. □

Si potrebbe pensare anche che una contraddizione si deriva da $A \wedge \neg B$, quindi $\neg(A \wedge \neg B)$; ma $\neg(A \wedge \neg B) \vdash A \to B$ e $A \to B \vdash \neg(A \wedge \neg B)$ (esercizio).

Il viceversa per esercizio.

Grazie alla validità delle leggi associative per unione e intersezione, queste operazioni possono essere generalizzate a più di due insiemi.

Se $A_1, \ldots, A_n$ sono n sottoinsiemi di U, la loro unione è l'insieme i cui elementi sono gli elementi di U che appartengono a qualche A_i, in simboli:

$$\bigcup_{i=1}^{n} A_i = \{x \in U \mid \text{ per qualche } i, 1 \leq i \leq n, x \in A_i\}$$

o anche

$$\bigcup_{i=1}^{n} A_i,$$

o semplicemente

$$\bigcup A_i.$$

L'intersezione generalizzata degli n insiemi è l'insieme degli elementi che appartengono a tutti gli A_i, in simboli:

$$\bigcap_{i=1}^{n} A_i = \{x \in U \mid \text{ per ogni } i, 1 \leq i \leq n, x \in A_i\}$$

o anche

$$\bigcap_{i=1}^{n} A_i,$$

o semplicemente

$$\bigcap A_i.$$

Si nota che la definizione dell'operazione generalizzata differisce da quella della operazione a due argomenti per l'uso dei quantificatori invece dei connettivi, rispettivamente $\exists$ per $\vee$ e $\forall$ per $\wedge$. Anche i quantificatori, si possono interpretare come connettivi generalizzati.

Per le operazioni generalizzate valgono molte delle leggi dell'unione e intersezione, opportunamente riformulate, ad esempio le proprietà commutativa, associativa e di assorbimento; valgono le leggi di De Morgan:

$$\sim(\bigcap_{i=1}^{n} A_i) = \bigcup_{i=1}^{n}(\sim A_i)$$

e

$$\sim(\bigcup_{i=1}^{n} A_i) = \bigcap_{i=1}^{n}(\sim A_i).$$

Valgono le leggi distributive di una operazione generalizzata rispetto a una normale (non con entrambe generalizzate):

$$(\bigcap_{i=1}^{n} A_i) \cup B = \bigcap_{i=1}^{n}(A_i \cup B)$$

e

$$(\bigcup_{i=1}^{n} A_i) \cap B = \bigcup_{i=1}^{n}(A_i \cap B).$$

Più in generale ancora, si definisce l'unione

$$\bigcup_{i \in I} A_i \quad \text{o} \quad \bigcup\{A_i \mid i \in I\}$$

per una famiglia di insiemi indiciata[7] da I ponendo che

$$x \in \bigcup_{i \in I} A_i \text{ se e solo se esiste un } i \in I \text{ per cui } x \in A_i,$$

e analogamente per l'intersezione. La definizione come si vede è la stessa, con "$i \in I$ al posto di "$1 \leq i \leq n$.

Esercizi

Esercizio 3.1.2. Completare la dimostrazione di tutte le leggi elencate.

Esercizio 3.1.3. Dimostrare che $X \triangle Y = \emptyset$ se e solo se $X = Y$. E se $X \triangle Y = U$?

Esercizio 3.1.4. Se R e S sono due simboli di predicato, la loro interpretazione in un universo U si indica con $R^U = \{x \in U \mid R(x) \in U\}$ e $S^U = \{x \in U \mid S(x) \in U\}$.

Se $R^U \subsetneq S^U$ quale è l'insieme definito in U da

$$R(x) \vee S(x) \to \forall x R(x)$$

e quale quello definito da

$$R(x) \vee S(x) \to \exists x S(x)?$$

3.2 Algebre di Boole

Dall'esempio svolto a pagina 79, relativo alla 5, si vede anche che in dimostrazioni di questo tipo fa comodo, per saltare qualche passaggio, fare appello ad altre delle leggi elencate — più semplici, o intuitive o semplicemente già dimostrate. Più in generale, una volta dimostrate alcune delle suddette leggi in modo diretto, è possibile derivare le altre in stile *algebrico*, usando quelle già dimostrate e le leggi dell'uguaglianza.

[7]Si chiama così e si indica anche con $\{A_i\}_{i \in I}$ un insieme i cui elementi corrispondono ciascuno ad un elemento di un insieme I, detto insieme degli indici.

Con leggi dell'uguaglianza si intendono le proprietà riflessiva, simmetrica e transitiva di =, rappresentate dalle formule

$$x = x$$
$$x = y \to y = x$$
$$x = y \land y = z \to x = z,$$

e le proprietà di sostituzione, che sono di due tipi:

$$t_1 = t_2 \to t[t_1] = t[t_2],$$

dove t_1 ed t_2 sono termini del linguaggio in uso, e $t[x]$ un altro termine contenente una variabile x ($t[t_i]$ si ottiene da $t[x]$ sostituendo t_i a x), e

$$t_1 = t_2 \to (A[t_1] \leftrightarrow A[t_2]),$$

dove $A[x]$ sta per una formula qualunque.

Queste leggi sono tacitamente usate nei passaggi di trasformazione di formule algebriche, o di proposizioni di qualunque linguaggio che contenga l'uguaglianza. I passaggi da un'uguaglianza ad un'altra presuppongono il *modus ponens*: da $t_1 = t_2$ a $t[t_1] = t[t_2]$ grazie a $t_1 = t_2 \to t[t_1] = t[t_2]$ e da $t_1 = t_2$ e $A[t_1]$ a $A[t_2]$ grazie a $t_1 = t_2 \to (A[t_1] \leftrightarrow A[t_2])$.

Nel considerare le leggi dell'algebra degli insiemi scritte nella forma di uguaglianze, e la loro derivazione algebrica da altre uguaglianze si esegue di fatto un cambiamento di linguaggio e di logica: si ha ora solo la relazione = e simboli di operazione con le quali si costruiscono termini a partire dai simboli $X, Y, \ldots$ che ora sono intesi come variabili ($\subseteq$ può essere introdotta per definizione); le formula sono quindi solo formule atomiche. $X, Y, \ldots$ si possono legittimamente considerare variabili in questo linguaggio, perché non compaiono più le $x, y, \ldots$, e le variabili variano sull'insieme dei sottoinsiemi di U.

Esempi

1. La 15 segue dalla 14 e dalla 11 con i passaggi

$$\sim\emptyset = U$$
$$\sim(\sim\emptyset) = \sim U$$
$$\emptyset = \sim U$$
$$\sim U = \emptyset.$$

2. La 17 segue dalla 16 e dalle 12, 14, 11, 4, nell'ordine, con i seguenti passaggi

$$\begin{aligned}
X \cap (\sim X) &= \emptyset \\
\sim(X \cap (\sim X)) &= \sim \emptyset \\
(\sim X) \cup (\sim(\sim X)) &= U \\
(\sim X) \cup X &= U \\
X \cup (\sim X) &= U.
\end{aligned}$$

3. La 18 segue da 17, 7, 1, 16 e 21 con i seguenti passaggi

$$\begin{aligned}
X \cup (\sim X) &= U \\
U &= X \cup (\sim X) \\
X \cap U &= X \cap (X \cup (\sim X)) \\
X \cap U &= (X \cap X) \cup (X \cap (\sim X)) \\
X \cap U &= X \cup \emptyset \\
X \cap U &= X.
\end{aligned}$$

4. La 31: $X \subseteq Y$ se e solo se $\sim X \cup Y = U$, segue dalla 30 e da De Morgan con 11 e 14.

Indagando la reciproca derivabilità delle varie leggi, ci si accorge che tutte (sia quelle elencate che altre, quelle che sono valide per ogni famiglia di sottoinsiemi di un insieme) sono derivabili dalle seguenti (conservando la numerazione di 4.1):

3	$X \cap Y = Y \cap X$	commutatività dell'intersezione
4	$X \cup Y = Y \cup X$	commutatività dell'unione
5	$X \cap (Y \cap Z) = (X \cap Y) \cap Z$	associatività dell'intersezione
6	$X \cup (Y \cup Z) = (X \cup Y) \cup Z$	associatività dell'unione
7	$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$	distributività di $\cap$ rispetto a $\cup$
8	$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$	distributività di $\cup$ rispetto a $\cap$
16	$X \cap \sim X = \emptyset$	legge dell'inverso per $\cap$
17	$X \cup \sim X = U$	legge dell'inverso per $\cup$
18	$X \cap U = X$	legge dell'elemento neutro per $\cap$
21	$X \cup \emptyset = X$	legge dell'elemento neutro per $\cup$.

Queste leggi si chiamano *assiomi delle algebre di Boole*. La scelta degli assiomi non è arbitraria (ci sono ragioni di analogia con altri sistemi di assiomi per altre strutture) ma non è univoca. Abbiamo visto ad esempio che se ci fosse la 1, la 18 sarebbe superflua. L'importante è la mutua e varia interderivabilità delle leggi tra loro, e che tutte le leggi valide per i sottoinsiemi di un insieme non vuoto U siano derivabili da quelle scelte come assiomi. La raccolta di queste negli assiomi è solo, inizialmente, una comodità mnemonica.

L'insieme dei sottoinsiemi di un insieme non vuoto U, con le operazioni $\sim, \cap, \cup$ e gli elementi speciali $\emptyset$ e U è un particolare esempio di algebra di Boole, che si chiama *algebra di insiemi*; ne vedremo altre.

Vediamo come si derivano dagli assiomi alcune delle altre leggi prima elencate.

1 $X = X \cap X$

$$\begin{array}{ll}
X = X \cap U & \text{per la 18} \\
X = X \cap (X \cup \sim X) & \text{per la 17} \\
X = (X \cap X) \cup (X \cap \sim X) & \text{per la 7} \\
X = (X \cap X) \cup \emptyset & \text{per la 16} \\
X = X \cap X & \text{per la 21.}
\end{array}$$

2 $X = X \cup X$ (esercizio)

20 $X \cap \emptyset = \emptyset$

$$\begin{array}{ll}
X \cap \emptyset = X \cap (X \cap \sim X) & \text{per la 16} \\
X \cap \emptyset = (X \cap X) \cap \sim X & \text{per la 5} \\
X \cap \emptyset = X \cap \sim X & \text{per la 1} \\
X \cap \emptyset = \emptyset & \text{per la 16.}
\end{array}$$

19 $X \cup U = U$ (esercizio).

Prima di considerare altre leggi, occorre dimostrare l'*unicità* degli elementi neutri e del complemento. Per quello dell'intersezione, questo significa:

34 Se $X \cap Y = Y$ per ogni Y, allora $X = U$.

Dimostrazione. Sostituendo U a Y si ha $X \cap U = U$ ma $X \cap U = X$ per la 18, quindi $X = U$. □

Per l'elemento neutro dell'unione, l'unicità significa:

35 Se $X \cup Y = Y$ per ogni Y, allora $X = \emptyset$ (esercizio).

L'unicità del complemento, o dell'inverso, è la proprietà che:

36 Se $X \cap Y = \emptyset$ e $X \cup Y = U$ allora $X = \sim Y$.

Dimostrazione.

$$
\begin{aligned}
X &= X \cap U && \text{per la 18}\\
&= X \cap (Y \cup \sim Y) && \text{per la 17}\\
&= (X \cap Y) \cup (X \cap \sim Y) && \text{per la 7}\\
&= \emptyset \cup (X \cap \sim Y) && \text{per l'ipotesi}\\
&= (Y \cap \sim Y) \cup (X \cap \sim Y) && \text{per la 16}\\
&= (Y \cup X) \cap \sim Y && \text{per la 7}\\
&= U \cap \sim Y && \text{per l'ipotesi}\\
&= \sim Y && \text{per la 18,}
\end{aligned}
$$

usando anche la 3. □

11 $X = \sim\sim X$

Dimostrazione. Siccome $X \cap \sim X = \emptyset$ e $X \cup \sim X = U$, per la 36 ora vista con $\sim X$ al posto di Y da 16 e 17 si ha $X = \sim\sim X$. □

13 $\sim(X \cup Y) = \sim X \cap \sim Y$

Dimostrazione. Per applicare la 36, facciamo vedere che

$$(\sim X \cap \sim Y) \cup (X \cup Y) = U$$

e

$$(\sim X \cap \sim Y) \cap (X \cup Y) = \emptyset.$$

La prima segue da questi passaggi (abbreviati, esplicitarli tutti per esercizio, serve anche la 19 di sopra):

$$
\begin{aligned}
(\sim X \cap \sim Y) \cup (X \cup Y) &= (\sim X \cup X \cup Y) \cap (\sim Y \cup X \cup Y)\\
&= U \cap U = U
\end{aligned}
$$

e la seconda (utilizzando 20) da:

$$
\begin{aligned}
(\sim X \cap \sim Y) \cap (X \cup Y) &= (\sim X \cap \sim Y \cap X) \cup (\sim X \cap \sim Y \cap Y)\\
&= \emptyset \cup \emptyset = \emptyset
\end{aligned}
$$

□

37 $X \cap Y = X$ se e solo se $X \cap \sim Y = \emptyset$

Dimostrazione.

$$\begin{aligned} X &= X \cap U && \text{per la 18} \\ &= X \cap (Y \cup \sim Y) && \text{per la 17} \\ &= (X \cap Y) \cup (X \cap \sim Y) && \text{per la 5} \end{aligned}$$

quindi se $X \cap \sim Y = \emptyset$ allora $X = X \cap Y$.

Viceversa se $X = X \cap Y$

$$\begin{aligned} U &= \sim X \cup X && \text{per la 17} \\ &= \sim X \cup (X \cap Y) \\ &= (\sim X \cup X) \cap (\sim X \cup Y) && \text{per la 6} \\ &= U \cap (\sim X \cup Y) && \text{per la 17} \\ &= \sim X \cup Y && \text{per la 18} \end{aligned}$$

Quindi $\emptyset = X \cap \sim Y$ per la 13 e la 15 (esercizio). □

9 $X \cap (X \cup Y) = X$

Dimostrazione. Si noti che

$$X \cap (X \cup Y) = (X \cap X) \cup (X \cap Y) = X \cup (X \cap Y)$$

per la 1 e la 5, per cui la 9 e la 10 si dimostrano insieme.

Per la 37 e la 13

$$X \cap (X \cup Y) = X \text{ se e solo se } X \cap (\sim X \cap \sim Y) = \emptyset$$

ma $X \cap (\sim X \cap \sim Y) = (X \cap \sim X) \cap \sim Y = \emptyset \cap \sim Y = \emptyset$, per l'associatività, la commutatività e la 20. □

Esercizi

Esercizio 3.2.1. Dimostrare:

i $A \cap (B \cup (C \setminus A)) = A \cap B$

ii $A \cap B \cap (A \cup B) = A \cap B$

iii $A \cup (C \cap (A \cup B)) = A \cup (C \cap B)$

iv $(A \setminus B) \cup (B \cap A) = A$

v $(A \cap (B \cup C)) \cap (\sim B \cup A) = (A \cap B) \cup (A \cap C)$.

Esercizio 3.2.2. Dimostrare le proprietà 22–33 della relazione $\subseteq$, a partire dagli assiomi, usando 28 come definizione di $\subseteq$[8].

Lo stesso, usando una volta 29, una volta 30 e una volta 31 come definizione di $\subseteq$.

Esercizio 3.2.3. Dimostrare, a partire dagli assiomi delle algebre di Boole, tutte le altre leggi sopra elencate per le operazioni di un'algebra di insiemi.

3.2.1 Algebra 2

Due altre notevoli algebre di Boole sono importanti, l'algebra **2** e l'algebra delle proposizioni.

Quando si dice che gli assiomi sopra elencati sono gli assiomi delle algebre di Boole, non si intende che i simboli di operazioni usati nella formulazione degli assiomi denotino le operazioni insiemistiche di unione, intersezione e complemento; altrimenti le uniche algebre di Boole sarebbero le algebre di insiemi. S'intende solo che siano operazioni rispettivamente binarie (le prime due) e unaria (la terza), e che soddisfino le proprietà espresse dagli assiomi per tutti gli elementi di un insieme non vuoto, che è l'universo della struttura.

Può essere utile addirittura riscrivere gli assiomi con altri simboli che non abbiamo un significato già consolidato[9]:

$x \circ y = y \circ x$	commutatività
$x + y = y + x$	commutatività
$x \circ (y \circ z) = (x \circ y) \circ z$	associatività
$x + (y + z) = (x + y) + z$	associatività
$x \circ (y + z) = (x \circ y) + (x \circ z)$	distributività
$x + (y \circ z) = (x + y) \circ (x + z)$	distributività
$x \circ (-x) = 0$	inverso
$x + (-x) = 1$	inverso
$x \circ 1 = x$	elemento neutro
$x + 0 = x$	elemento neutro

[8] La 22 e la 27, insieme a "$X = Y$ se e solo se $X \subseteq Y$ e $Y \subseteq X$" stabiliscono che $\subseteq$ è una relazione di *ordine parziale*.

[9] Con l'ordine di priorità $-$, $\circ$, $+$. $\circ$ è una pallina e non un punto, per distinguerla dalla moltiplicazione numerica.

Si indichi con $\leq$ la relazione definita da $x \leq y \leftrightarrow x \circ y = x$ (o da $x \leq y \leftrightarrow x \circ (-y) = 0$ o ancora da $x \leq y \leftrightarrow x + y = y$; a seconda dei casi converrà usare l'una o l'altra definizione).

La relazione $\leq$ è un ordine parziale per l'esercizio 2 di 4.2.1.

Si ha $0 \leq x \leq 1$ per ogni x (esercizio). Questo significa che 0 è il minimo rispetto alla relazione $\leq$ e 1 il massimo.

Inseriamo qui una dimostrazione dell'equivalenza tra le due definizioni di $\leq$ mediante $\circ$, dove si noterà l'analogia formale con quella fatta per la legge 37 dell'algebra degli insiemi.

Se

$$x \circ y = x$$

allora

$$\begin{aligned}
1 &= x + (-x)\\
1 &= (x \circ y) + (-x)\\
1 &= (x + (-x)) \circ (y + (-x))\\
1 &= y + (-x)\\
0 &= x \circ (-y).
\end{aligned}$$

Viceversa se

$$x \circ (-y) = 0$$

allora

$$\begin{aligned}
x &= x \circ 1\\
x &= x \circ (y + (-y))\\
x &= (x \circ y) + (x \circ (-y))\\
x &= x \circ y.
\end{aligned}$$

Esercizio 3.2.4. Dimostrare l'equivalenza con la (versione corrispondente della) 29 di pagina 79.

Si dimostra (esercizio) $x + y = \sup\{x, y\}$ e $x \circ y = \inf\{x, y\}$, sup e inf rispetto a $\leq$: significano che $x + y$ è il minimo (rispetto a $\leq$) elemento maggiore o uguale sia di x che di y, e che $x \circ y$ è il massimo elemento minore o uguale sia di x che di y.

Si può anche definire un'algebra di Boole come un insieme parzialmente ordinato da una relazione $\leq$ tale che

1. per ogni x, y esistano $\sup\{x, y\}$ e $\inf\{x, y\}$, che sono denotati rispettivamente $x + y$ e $x \circ y$,

2. con un minimo 0 e un massimo 1, e per ogni x un elemento $-x$ tale che $\sup\{x, -x\} = 1$ e $\inf\{x, -x\} = 0$, e

3. valgano le proprietà distributive.

L'algebra **2** è l'algebra il cui universo è $\{0, 1\}$ con $0 < 1$, rappresentata dal diagramma

$$\begin{array}{c} 1 \\ \uparrow \\ 0 \end{array}$$

dove $\uparrow$ è $<$ e in questo caso $x + y = \max\{x, y\}$ e $x \circ y = \min\{x, y\}$.

L'algebra **2** è l'algebra dei valori di verità. Le sue tre operazioni sono quelle che intervengono come abbiamo visto nel calcolo dei valori di verità di negazioni, disgiunzioni e congiunzioni.

Esistono altre algebre di Boole finite, come ad esempio l'algebra **4**

$$\begin{array}{ccccc} & & 1 & & \\ & \nearrow & & \nwarrow & \\ a & & & & b \\ & \nwarrow & & \nearrow & \\ & & 0 & & \end{array}$$

dove a e b sono inconfrontabili rispetto a $\leq$; $\leq$ è proprio parziale.

Esercizio 3.2.5. Definire le operazioni in modo che questa struttura ordinata diventi un'algebra di Boole.

Esercizio 3.2.6. Dimostrare che è l'algebra dei sottoinsiemi di un universo con due elementi.

3.2.2 Algebra delle proposizioni

L'*algebra delle proposizioni* si ottiene nel seguente modo; già si sono dimostrate (considerando anche gli esercizi) quasi tutte le leggi logiche che hanno lo stesso nome degli assiomi delle algebre di Boole:

$$\begin{array}{ll}
A \land B \leftrightarrow B \land A & \text{commutatività} \\
A \lor B \leftrightarrow B \lor A & \text{commutatività} \\
A \land (B \land C) \leftrightarrow (A \land B) \land C & \text{associatività} \\
A \lor (B \lor C) \leftrightarrow (A \lor B) \lor C & \text{associatività} \\
A \land (B \lor C) \leftrightarrow (A \land B) \lor (A \land C) & \text{distributività} \\
A \lor (B \land C) \leftrightarrow (A \lor B) \land (A \lor C) & \text{distributività.}
\end{array}$$

Le equivalenze non sono uguaglianze ma si possono trasformare in vere uguaglianze tra (nuovi) oggetti con la seguente costruzione (che è frequente in matematica[10] e si chiama "passaggio al quoziente rispetto a una relazione di equivalenza").

La relazione $\equiv$ è una *relazione di equivalenza*, vale a dire soddisfa le proprietà:

$$\begin{array}{lr} A \equiv A & \text{riflessiva} \\ \text{se } A \equiv B \text{ allora } B \equiv A & \text{simmetrica} \\ \text{se } A \equiv B \text{ e } B \equiv C \text{ allora } A \equiv C & \text{transitiva.} \end{array}$$

Si definisce allora per ogni A la *classe di equivalenza* di A come

$$[A] = \{B | A \equiv B\}$$

e si ha che

$$[A] = [B] \text{ se e solo se } A \equiv B$$

(esercizio).

Date due proposizioni A e B, esse o sono logicamente equivalenti o no. Nel primo caso, come si è detto, $[A] = [B]$. Nel secondo caso le due classi $[A]$ e $[B]$ sono disgiunte: se infatti ci fosse un elemento C in comune, vorrebbe dire che $A \equiv C$ e che $B \equiv C$, ma allora per la transitività si avrebbe $A \equiv B$ e $[A] = [B]$.

A si dice un rappresentante della classe $[A]$; ogni classe ha più rappresentanti, anzi infiniti. Se $B \in [A]$ allora $B \equiv A$ quindi $[A] = [B]$ e B è un altro rappresentante di $[A]$. In particolare ad esempio $[A] = [A \wedge A] = [A \wedge A \wedge A] \dots$

Si possono definire tra queste classi le seguenti operazioni:

$$\begin{aligned} -[A] &= [\neg A] \\ [A] \circ [B] &= [A \wedge B] \\ [A] + [B] &= [A \vee B]. \end{aligned}$$

Si vede il vantaggio di aver formulato gli assiomi delle algebre di Boole in un linguaggio algebrico astratto, perché ora si possono usare questi simboli. Non si potrebbe usare $\neg$ per il complemento, perché $\neg$ ha già un significato, per il quale $\neg[A]$ non ha senso, né $\sim$, perché $\sim[A] \neq [\neg A]$. Lo stesso per $+$ e $\circ$.

Le definizioni sono ben poste, in questo senso. Si tratta di operazioni sulle classi, ma la loro definizione fa riferimento ad un particolare rappresentante

[10] Ad esempio in geometria la direzione di una retta è, nella terminologia introdotta qui di seguito, la classe di equivalenza delle rette parallele a quella data.

delle classi. Ad esempio $-[A]$ è definita con $\neg A$ e non ad esempio con $\neg\neg\neg A$. Se si cambia il rappresentante di una classe, si vuole che il risultato, che è una classe, sia lo stesso.

In effetti è così per le operazioni sopra definite. Ad esempio se $A_1 \equiv A$ e $B_1 \equiv B$, siccome $A_1 \wedge B_1 \equiv A \wedge B$ (esercizio — si veda anche alla fine di questo paragrafo) si ha $[A_1] \circ [B_1] = [A \wedge B]$, così come $[A] \circ [B] = [A \wedge B]$, quindi $[A_1] \circ [B_1] = [A] \circ [B]$.

Si giustifica in questo modo la dizione "a meno di equivalenza" con cui una proposizione è considerata uguale ad ogni altra ad essa logicamente equivalente, o almeno indistinguibile da quelle, ai fini della trattazione semantica.

Date queste definizioni, le precedenti equivalenze danno allora origine alle uguaglianze:

$$\begin{array}{rl}
[A] \circ [B] = [B] \circ [A] & \text{commutatività di } \circ \\
[A] + [B] = [B] + [A] & \text{commutatività di } + \\
[A] \circ ([B] \circ [C]) = ([A] \circ [B]) \circ [C] & \text{associatività di } \circ \\
[A] + ([B] + [C]) = ([A] + [B]) + [C] & \text{associatività di } + \\
[A] \circ ([B] + [C]) = ([A] \circ [B]) + ([A] \circ [C]) & \text{distributività} \\
[A] + ([B] \circ [C]) = ([A] + [B]) \circ ([A] + [C]) & \text{distributività.}
\end{array}$$

Tutte le tautologie sono tra loro equivalenti, e non equivalenti a nessuna proposizione non logicamente valida; lo stesso per le contraddizioni; denotiamo con 1 la classe delle tautologie, e con 0 la classe delle contraddizioni.

Allora $[A] \circ (-[A]) = [A \wedge \neg A] = 0$ e $[A] + (-[A]) = [A \vee \neg A] = 1$ e possiamo quindi aggiungere:

$$\begin{array}{rl}
[A] \circ (-[A]) = 0 & \text{inverso} \\
[A] + (-[A]) = 1 & \text{inverso} \\
[A] \circ 1 = [A] & \text{elemento neutro} \\
[A] + 0 = [A] & \text{elemento neutro}
\end{array}$$

completando la lista degli assiomi delle algebre di Boole.

Le ultime due leggi seguono dal fatto (o lo esprimono in altra forma) che se T è una tautologia $A \wedge T \equiv A$ e se F è una contraddizione allora $A \vee F \equiv A$.

La relazione $[A] \leq [B]$ è definita da $[A] \circ [B] = [A]$, oppure dall'equivalente $[A] \circ -[B] = 0$, o ancora dalle altre condizioni equivalenti (esercizio 3 di 4.2.1).

Dall'equivalenza booleana delle diverse definizioni di $\leq$ si deriva la seguente proprietà logica, che

$$A \equiv A \wedge B \text{ se e solo se } \models A \to B.$$

Una dimostrazione logica di questo fatto ricalca la dimostrazione algebrica di sopra.

La seguente è una deduzione del fatto che $\models A \to B$ segue da $A \equiv A \wedge B$:

$$\begin{gathered} A \vee \neg A \\ (A \wedge B) \vee \neg A \\ (A \vee \neg A) \wedge (B \vee \neg A) \\ B \vee \neg A \\ A \to B. \end{gathered}$$

Ogni proposizione di questa lista o è una tautologia o segue logicamente dalle precedenti e da $A \equiv A \wedge B$, quindi l'ultima è una tautologia.

Viceversa, se $\models A \to B$, quindi $A \wedge \neg B$ è una contraddizione,

$$\begin{gathered} A \leftrightarrow A \wedge (B \vee \neg B) \\ A \leftrightarrow (A \wedge B) \vee (A \wedge \neg B) \\ A \leftrightarrow A \wedge B. \end{gathered}$$

La corrispondenza tra le deduzioni algebriche e quelle logiche è fondata sulla corrispondenza tra $[A] \leq [B]$ e $\models A \to B$.

Il fatto che per ogni A, $0 \leq [A] \leq 1$ corrisponde al fatto che una contraddizione implica qualsiasi proposizione, e una tautologia è implicata da qualsiasi proposizione.

La relazione booleana $\leq$ ha le seguenti proprietà, che se $a \leq b$ allora $-b \leq -a$ e per ogni c, $c \circ a \leq c \circ b$ e $c + a \leq c + b$ (esercizio).

Queste proprietà corrispondono per l'implicazione al fatto che se $\models A \to B$ allora $\models \neg B \to \neg A$ e per ogni C, $\models C \wedge A \to C \wedge B$ e $\models C \vee A \to C \vee B$ (vedi esercizi).

La proprietà transitiva di $\leq$ corrisponde alla transitività del condizionale, mentre la proprietà di sostituzione $t_1 = t_2 \to t[t_1] = t[t_2]$ corrisponde ad un'analoga proprietà logica: se in una proposizione si sostituisce una sottoproposizione con una equivalente, il risultato è una proposizione equivalente a quella iniziale.

Conviene indicare l'operazione di *rimpiazzamento* di una sottoproposizione B con C in una proposizione A, con la notazione: $A[B /\!\!/ C]$.

Si ha allora che

$$\text{se } B \equiv C \text{ allora } A \equiv A[B /\!\!/ C].$$

Nell'esempio di sopra $A \vee \neg A \equiv (A \wedge B) \vee \neg A$ poiché $A \equiv A \wedge B$.

Nella dimostrazione, per trattare i vari casi, si fa uso dei seguenti fatti: per ogni A e B,

$$\text{se } A \equiv B, \text{ allora } \neg A \equiv \neg B$$
$$\text{se } A_1 \equiv A_2 \text{ e } B_1 \equiv B_2, \text{ allora } A_1 \bullet B_1 \equiv A_2 \bullet B_2$$

che si dimostrano facilmente con le tavole di verità per i vari connettivi.

Esercizi

Esercizio 3.2.7. Spiegare perché $\sim[A] \neq [\neg A]$, in generale.

Esercizio 3.2.8. Verificare che $[A] \circ [B] = [A] \cap [B]$, mentre $[A] + [B] \neq [A] \cup [B]$ (trovare un controesempio).

Esercizio 3.2.9. Dimostrare che $A \equiv A \vee B$ se e solo se $\models B \to A$.

Esercizio 3.2.10. Dimostrare algebricamente che se $\models A \to B$ allora $\models \neg B \to \neg A$ e per ogni C, $\models C \wedge A \to C \wedge B$ e $\models C \vee A \to C \vee B$.

Capitolo 4

Forme normali

Dopo aver imparato le definizioni riguardanti la semantica delle proposizioni, e alcune prime tecniche per stabilire in particolare se sono tautologie, sia direttamente con il calcolo del valori di verità sia deducendole da altre con passaggi logici o algebrici booleani, passiamo a porci alcuni problemi metateorici sul linguaggio proposizionale.

4.1 Definibilità dei connettivi

Ad ogni proposizione è associata una tavola di verità , come abbiamo visto negli esempi della sezione 2.2. Viceversa, data una qualunque tavola di verità, come ad esempio

p	q	r	?
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

esiste una proposizione scritta utilizzando soltanto i connettivi $\neg$, $\wedge$, $\vee$ che ha quella data come sua tavola di verità associata.

La proposizione si costruisce nel seguente modo, appoggiandosi come esempio alla tavola di sopra. Sarà una disgiunzione con tanti disgiunti quante sono nella tavola le righe che hanno il valore 1, quindi $A_1 \vee A_2 \vee A_3 \vee A_4$; ogni disgiunto A_i dovrà essere vero solo per l'interpretazione della riga corrispondente; la riga assegna valori 0, 1 alle lettere, quindi 1 a certe lettere e 1 alle negazioni di certe altre lettere; una congiunzione è vera se e solo se tutti i congiunti sono veri; A_i potrà quindi essere una congiunzione di tante proposizioni quante sono le colonne di entrata della tavola, nell'esempio 3, e ciascuna di queste proposizioni sarà una lettera o la negazione di quella lettera a seconda che nella riga corrispondente la lettera abbia il valore 1 oppure 0. Quindi

$$(\neg p \land \neg q \land \neg r) \lor (\neg p \land \neg q \land r) \lor (\neg p \land q \land \neg r) \lor (\neg p \land q \land r).$$

Per le proprietà della valutazione della disgiunzione e congiunzione — che una disgiunzione è vera se e solo se almeno un disgiunto è vero, e una congiunzione se e solo se tutti i congiunti sono veri — e della negazione, si può facilmente vedere procedendo al contrario che la tavola associata a questa proposizione è uguale alla tavola data, che era la tavola di $p \lor r \to \neg p \land (q \to r)$.

Il risultato si esprime anche dicendo che tutte le funzioni di verità sono *definibili* in termini dell'insieme di connettivi $\{\neg, \land, \lor\}$, o che questo è un insieme *adeguato* di connettivi.Questo significa che non si è perso nulla, quanto a capacità espressiva, non ammettendo nell'alfabeto altri connettivi, ad esempio quello per la duplice negazione "né ... né"; se avessimo introdotto un connettivo $\uparrow$ o NOR per questa combinazione di proposizioni, con la tavola

A	B	$A \uparrow B$
0	0	1
0	1	0
1	0	0
1	1	0

a posteriori potremmo ora sostituire ogni occorrenza della proposizione $p \uparrow q$ con l'equivalente $\neg p \land \neg q$[1].

Si faccia attenzione a cosa significa che un *simbolo* è *definibile*, a differenza ad esempio dalla definibilità di un insieme.

Un insieme $X \subseteq U$ è definibile in U se X è l'insieme di verità in U di una formula.

Un simbolo di operatore binario $\bullet$ si dice definibile (in termini di altri) se $p \bullet q \leftrightarrow A[p, q]$ oppure $p \bullet q = A[p, q]$, a seconda che $p \bullet q$ sia una formula oppure un termine, dove A è un'espressione che non contiene $\bullet$ e contiene solo gli altri simboli o nozioni nei termini dei quali $\bullet$ si dice definito.

S'intende che il bicondizionale o l'uguaglianza definitorie devono essere valide nel contesto in esame: in logica sarà $\models p \bullet q \leftrightarrow A[p, q]$, mentre

[1]Se c'è una sola riga con valore 1, la proposizione costruita come detto sopra è della forma A_1, dove A_1 è una congiunzione. Si può dire tuttavia che anche in questo caso la proposizione associata alla tavola è una disgiunzione, pensando che $A_1 \equiv A_1 \lor A_1$.

una uguaglianza $p \bullet q = A[p, q]$ deve essere dimostrata nella relativa teoria, aritmetica o algebra o altro.

Analogamente se il numero di argomenti è diverso da 2.

Ad esempio in geometria piana per due rette "$r \parallel s \leftrightarrow r$ e s non si intersecano" ($r \parallel s \leftrightarrow \neg\exists x(Q(x, r) \wedge Q(x, s))$), in aritmetica "$x \mid y \leftrightarrow$ esiste uno z per cui $x \cdot z = y$", o il simbolo di potenza al quadrato "$x^2 = x \cdot x$", nell'algebra degli insiemi "$X \setminus Y = X \cap \sim Y$".

Ma il precedente risultato dice anche che gli stessi connettivi del linguaggio proposizionale sono sovrabbondanti, perché $\{\neg, \wedge, \vee\}$ è adeguato, e neanche il più ridotto possibile. Quando un sistema adeguato è minimale, nel senso che nessun suo sottoinsieme proprio è ancora adeguato, si chiama una *base*, in analogia con le basi degli spazi vettoriali (si vedano gli esercizi).

Si ha che $p \oplus q$ risulta equivalente a $(\neg p \wedge q) \vee (p \wedge \neg q)$, e $p \to q \equiv (\neg p \wedge \neg q) \vee (\neg p \wedge q) \vee (p \wedge q)$ e analogamente $p \leftrightarrow q$ (esercizio).

Ognuna di queste equivalenze comporta l'eliminabilità del connettivo definito, cioè che all'interno di una proposizione una sottoproposizione, ad esempio della forma $A \to B$, può essere rimpiazzata dalla proposizione equivalente $(\neg A \wedge \neg B) \vee (\neg A \wedge B) \vee (A \wedge B)$.

Esercizi

Esercizio 4.1.1. Dimostrare che $\{\neg, \wedge\}$ e $\{\neg, \vee\}$ sono due basi di connettivi, definendo la disgiunzione nel primo e la congiunzione nel secondo.

Esercizio 4.1.2. Dimostrare che $\{\neg, \to\}$ è una base di connettivi.

Esercizio 4.1.3. Dimostrare che il connettivo "né ... né" da solo costituisce una base, definendo in termini di esso la negazione e la congiunzione.

Esercizio 4.1.4. Scrivere la funzione di verità del connettivo $\downarrow$ o NAND, "non entrambe", o "non sia ... sia", e dimostrare che costituisce da solo una base di connettivi.

Esercizio 4.1.5. Definire in termini di $\neg, \wedge$ e $\vee$ il connettivo $m(p, q, r) = 1$ se e solo se la maggioranza delgi argomenti vale 0.

Esercizio 4.1.6. Esaminare tutte le tavole di verità a una entrata, e spiegare perché non esiste un connettivo per "è necessario che".

Esercizio 4.1.7. Discutere se è possibile ripetere la trattazione di questo paragrafo con $\oplus$ al posto di $\vee$ (cioè associare a ogni tavola una proposizione con $\neg, \wedge, \oplus$ che abbia quella data come sua tavola di verità).

L'insieme $\{\neg, \wedge, \oplus\}$ è adeguato? E $\{\neg, \oplus\}$? E $\{\oplus, \wedge\}$?

Esercizio 4.1.8. Dimostrare che $\{\oplus, \rightarrow\}$ è una base, definendo $\neg$.

Esercizio 4.1.9. Dimostrare che $\{\rightarrow\}$ non è adeguato.

Esercizio 4.1.10. In alcune presentazioni dei linguaggi proposizionali l'alfabeto contiene un simbolo $\perp$, per il quale si conviene che $(\perp)$ è una proposizione e che in ogni intepretazione i, $i^*(\perp) = 0$.

Dimostrare che $\{\perp, \rightarrow\}$ è una base.

4.2 Forme normali disgiuntive

La proposizione costruita a partire da una tavola di verità nel modo sopra descritto ha una forma particolare. Si chiami *letterale* una proposizione che sia o una lettera p, letterale *positivo*, o la negazione di una lettera $\neg p$, letterale *negativo*.

La proposizione associata alla tavola ha dunque la forma di una disgiunzione di congiunzioni di letterali. Una tale forma di chiama *forma normale disgiuntiva*. Poiché è evidente che

Osservazione 4.2.1. Per ogni A e B che contengano le stesse lettere,

$$A \equiv B \text{ se e solo se } A \text{ e } B \text{ hanno la stessa tavola di verità}$$

si può concludere che

Teorema 4.2.2. *Per ogni proposizione A esiste una proposizione con le stesse lettere che è in forma normale disgiuntiva ed è logicamente equivalente ad A.*

Dimostrazione. Come nell'esempio di sopra, data A si calcoli la sua tavola, quindi si costruisca la proposizione in forma normale disgiuntiva associata alla tavola.

Nel caso che la tavola di A non abbia alcun 1 nella colonna dei valori, quindi che A sia una contraddizione, la proposizione equivalente in forma normale disgiuntiva si può scrivere nella forma $(\neg p \wedge p) \vee \ldots \vee (\neg q \wedge q)$ come disgiunzione di contraddizioni elementari, una per ogni lettera di A. □

Anche una proposizione come $\neg p \vee q$ è in forma normale disgiuntiva, perchè il concetto di congiunzione e disgiunzione è usato ovviamente in senso generalizzato, ammettendo due o più componenti, o anche una sola[2]. Le proposizioni in forma normale disgiuntiva associate a tavole di proposizioni non contradditorie hanno l'ulteriore proprietà che in ogni disgiunto compaiono

[2] Vedi anche la precedente nota 1.

le stesse lettere, e che in ogni congiunzione ogni lettera compare una sola volta, o positiva o negata[3]. Qualche volta si usa l'aggettivo *regolare* per indicare questa caratteristica delle forme normali. Una proposizione in forma normale disgiuntiva regolare permette di leggere direttamente i modelli della proposizioni, uno per ogni disgiunto:

$$(\neg p \land q) \lor (p \land \neg q)$$

ha due modelli, $i_1(p) = 0$ e $i_1(q) = 1$, e $i_2(p) = 1$ e $i_2(q) = 0$.

Tale possibilità di lettura dei modelli sussiste peraltro anche per le forme normali disgiuntive non regolari, considerando però le interpretazioni come definite in modo arbitrario sulle lettere che non occorrono in alcuni disgiunti:

$$(\neg p \land q) \lor p$$

ha tre modelli: da $\neg p \land q$ viene $i_1(p) = 0$ e $i_1(q) = 1$, e da p viene $i(p) = 1$, che però ne riassume due: $i_2(p) = 1$ e $i_2(q) = 1$, e $i_3(p) = 1$ e $i_3(q) = 0$.

Qualche volta, sempre per le forme non regolari, disgiunti diversi hanno modelli in comune; e ovviamente se in una congiunzione occorre sia una lettera sia la sua negazione quella congiunzione non ha modelli.

4.3 Forme normali congiuntive

Un altro modo di associare a una tavola una proposizione scritta solo con i connettivi $\neg$, $\land$ e $\lor$ è il seguente, dove sono scambiati i ruoli di 0 e 1 e di congiunzione e disgiunzione: si cerca ora una proposizione che sia falsa esattamente nei casi prescritti dalla tavola data. In riferimento allo stesso esempio di prima, la proposizione deve essere falsa solo ed esattamente in corrispondenza alle ultime quattro righe della tavola, sarà perciò una congiunzione $A_5 \land A_6 \land A_7 \land A_8$, e ogni A_i sarà la disgiunzione di tre letterali, ogni letterale positivo o negativo a seconda che nella riga in questione la lettera abbia il valore 0 oppure 1. Quindi:

$$(\neg p \lor q \lor r) \land (\neg p \lor q \lor \neg r) \land (\neg p \lor \neg q \lor r) \land (\neg p \lor \neg q \lor \neg r).$$

Per confermare che questa proposizione ha la tavola data come sua tavola di verità occorre questa volta ricordare che una congiunzione è Êfalsa se e solo se una delle proposizioni congiunte è falsa, e che una disgiunzione è falsa se e solo se tutte le proposizioni disgiunte sono false.

Una proposizione che sia una congiunzione di disgiunzioni di letterali si dice in *forma normale congiuntiva*.

[3]Questa disgiunzione nel testo è esclusiva.

Una disgiunzione di letterali (una forma che interviene in molte considerazioni) si chiama *clausola*.

Esempio 4.3.1. La forma normale congiuntiva di $p \to q$, applicando il procedimento descritto, è $\neg p \lor q$, che è forma congiuntiva, se si considera, come si considera, la congiunzione in senso generalizzato; $\neg p \lor q$ è dunque in forma sia congiuntiva sia disgiuntiva.

Come sopra, risolvendo a parte anche il caso in cui nella tavola non ci siano 0, si ha:

Teorema 4.3.2. *Per ogni proposizione A esiste una proposizione con le stesse lettere che è in forma normale congiuntiva ed è equivalente ad A.*

Le forme normali, non necessariamente regolari, sono convenienti per verificare in modo efficiente (alla sola scansione e ispezione della lista) la validità logica o l'insoddisfacibilità , ma ciascuna forma è adeguata solo per una delle due proprietà.

Teorema 4.3.3. *Una proposizione in forma normale congiuntiva è una tautologia se e solo se in ogni sua clausola c'è una lettera che occorre sia positiva sia negata.*

Una proposizione in forma normale disgiuntiva è insoddisfacibile se e solo se in ogni suo disgiunto c'è una lettera che occorre sia positiva sia negata.

Dimostrazione. Per le forme congiuntive, una clausola in cui occorra una lettera e la negazione della stessa lettera è una tautologia, e una congiunzione è una tautologia se e solo se lo sono le sue componenti. Una clausola in cui non si verifichi la presenza di una lettera e della sua negazione può assumere il valore 1 se a tutti i letterali si assegna il valore 1 interpretando a 1 le lettere dei letterali positivi e a 0 le lettere dei letterali negativi.

Un ragionamento analogo vale per le forme disgiuntive. □

Si noti che due proposizioni equivalenti non debbono necessariamente avere le stesse lettere, ad esempio $q \land (\neg p \lor p)$ è equivalente a q, e $\neg p \lor p$ è equivalente a $q \to q$ (sono tutt'e due tautologie); quando si controlla che per ogni interpretazione le due proposizioni hanno lo stesso valore si considerano interpretazioni definite sull'insieme più ampio di lettere, ma si possono trascurare in una proposizione i valori delle lettere non occorrenti.

Le proposizioni in forma normale che si ottengono da una tavola non sono sempre le più semplici possibili. Se ad esempio il criterio che interessa è quello della lunghezza, la forma $\neg p \lor q$ per il condizionale è preferibile alla forma normale disgiuntiva regolare che si ottiene dalla tavola di verità. A

$\neg p \lor q$ si può passare dalla forma normale disgiuntiva regolare $(\neg p \land \neg q) \lor (\neg p \land q) \lor (p \land q)$ con i seguenti passaggi:

$$
\begin{gathered}
(\neg p \land \neg q) \lor (\neg p \land q) \lor (p \land q) \\
(\neg p \land (\neg q \lor q)) \lor (p \land q) \\
\neg p \lor (p \land q) \\
(\neg p \lor p) \land (\neg p \lor q) \\
\neg p \lor q
\end{gathered}
$$

applicando le leggi distributive e la semplificazione delle tautologie (si noti che nel corso delle trasformazioni si passa anche per proposizioni non in forma normale).

Come mostra l'esempio, esistono quindi diverse forme normali disgiuntive (e lo stesso per le congiuntive) equivalenti a una data proposizione; si parlerà perciò solo impropriamente della forma normale disgiuntiva (o congiuntiva) di una proposizione A, ma si userà ugualmente tale dizione, intendendola a meno di equivalenza logica; si chiamerà in tal modo una qualunque forma normale disgiuntiva (o congiuntiva) che sia equivalente ad A[4], e si potrà anche scrivere, se conveniente, DNF(A) (rispettivamente CNF(A)).

Il risultato generale che ogni proposizione è equivalente a una proposizione in forma normale disgiuntiva o congiuntiva si può ottenere anche applicando un algoritmo FORMA NORMALE di trasformazioni successive come nell'esempio di sopra per il condizionale.

Il procedimento è il seguente:

- eliminare $\oplus$, $\leftrightarrow$ e $\rightarrow$
- spostare $\neg$ verso l'interno con le leggi di De Morgan
- cancellare le doppie negazioni, con la legge della doppia negazione
- cancellare le ripetizioni, con le leggi di idempotenza
- applicare ripetutamente le leggi distributive.

L'ultima indicazione può sembrare vaga, ma si può rendere più precisa e deterministica. Con i passi precedenti si è ottenuta una proposizione equivalente che è formata a partire da letterali con applicazioni ripetute di $\land$ e $\lor$, anche se non necessariamente nell'ordine che produce una forma normale. Supponiamo di volerla trasformare in forma normale congiuntiva (per la

[4]Non necessariamente con le stesse lettere, come mostra l'esempio delle due forme normali disgiuntive $p \lor (q \land \neg q) \equiv p$.

forma normale disgiuntiva il procedimento è lo stesso con scambiati i ruoli di $\wedge$ e $\vee$).

Consideriamo il connettivo principale della proposizione; se è $\wedge$, passiamo alle due sottoproposizioni immediate trasformandole separatamente con il procedimento sotto descritto[5] e facendo alla fine la congiunzione delle due forme congiuntive così ottenute; se è $\vee$, e la proposizione è della forma $A \vee B$, è necessaria qualche preparazione.

Se in A non occorresse per nulla $\wedge$, potremmo lavorare su B come detto sotto, dopo aver fatto, per la precisione, lo scambio con $B \vee A$. Possiamo allora supporre che A sia della forma $C \wedge D$, perché se A a sua volta fosse una disgiunzione $C \vee D$, potremmo considerare al suo posto l'equivalente $C \vee (D \vee B)$ e andare a cercare $\wedge$ in C, oppure in D dopo aver fatto lo scambio con l'equivalente $D \vee (C \vee B)$.

La proposizione data si trasforma allora nella equivalente $(C \vee B) \wedge (D \vee B)$ e possiamo applicare ricorsivamente e separatamente il procedimento alle due proposizioni più corte $C \vee B$ e $D \vee B$, e alla fine ricongiungere con $\wedge$ i due risultati.

Siccome ogni stadio del procedimento porta a ripartire col procedimento su proposizioni più corte, e la lunghezza è un numero naturale, il processo deve terminare, e a un certo punto si arriva a una $E \vee B$ dove in E non occorre più $\wedge$. Allora se in B non occorre $\wedge$ si ha una disgiunzione che fornisce uno dei congiunti della forma normale congiuntiva. Altrimenti B è della forma $F \wedge G$ quindi $E \vee B$ equivalente a $(E \vee F) \wedge (E \vee G)$ e si riapplica l'algoritmo a $E \vee F$ e $E \vee G$.

Esempio 4.3.4. Da

$$\begin{aligned}
&(p \to q) \to (r \vee \neg p) \\
&\neg(p \to q) \vee (r \vee \neg p) \\
&\neg(\neg p \vee q) \vee (r \vee \neg p) \\
&(\neg\neg p \wedge \neg q) \vee (r \vee \neg p) \\
&(p \wedge \neg q) \vee (r \vee \neg p),
\end{aligned}$$

che è in forma normale disgiuntiva

$$(p \wedge \neg q) \vee r \vee \neg p$$

con due disgiunti unitari r e $\neg p$. Se invece si vuole la forma normale congiuntiva, si continua con

$$\begin{gathered}
(p \vee (r \vee \neg p)) \wedge (\neg q \vee (r \vee \neg p)) \\
(p \vee r \vee \neg p) \wedge (\neg q \vee r \vee \neg p).
\end{gathered}$$

[5] L'algoritmo che stiamo presentando è un algoritmo ricorsivo.

Esempio 4.3.5. Trasformare la forma normale disgiuntiva $(p \wedge \neg q) \vee (\neg p \wedge q)$ in forma normale congiuntiva:

$$(p \wedge \neg q) \vee (\neg p \wedge q)$$
$$(p \vee (\neg p \wedge q)) \wedge (\neg q \vee (\neg p \wedge q)).$$

Il primo congiunto si trasforma in

$$(p \vee \neg p) \wedge (p \vee q),$$

il secondo in

$$(\neg q \vee \neg p) \wedge (\neg q \vee q),$$

quindi la proposizione in

$$(p \vee \neg p) \wedge (p \vee q) \wedge (\neg q \vee \neg p) \wedge (\neg q \vee q),$$

da cui si possono ancora eliminare le tautologie, ottenendo

$$(p \vee q) \wedge (\neg q \vee \neg p).$$

Non è detto che questo procedimento, che ha il merito di far vedere la terminazione del compito, se lo si segue come filo d'Arianna, sia sempre il più efficiente; può essere utilmente integrato con l'applicazione in itinere dell'eliminazione delle ripetizioni, e con l'eliminazione delle tautologie dalle congiunzioni, e della contraddizioni dalle disgiunzioni, *ogni volta* che sia possibile; sono utili le leggi di assorbimento ed equivalenze come $\neg(A \to B) \equiv A \wedge \neg B$; oppure ci sono scorciatoie come quando, volendo mirare a una forma congiuntiva, si incontra una sottoproposizione della forma $(A \wedge B) \vee (C \wedge B)$ che conviene rimpiazzare direttamente con $(A \vee C) \wedge B$.

Esempio 4.3.6. Trasformare $(\neg p \wedge q) \vee (\neg q \wedge \neg p)$ in forma normale congiuntiva.

Con la distributività (e la commutatività) si ottiene subito

$$\neg p \wedge (q \vee \neg q)$$

e quindi $\neg p$. Applicando invece l'algoritmo FORMA NORMALE si ottiene

$$((\neg p \wedge q) \vee \neg q) \wedge ((\neg p \wedge q) \vee \neg p).$$

Il secondo congiunto è equivalente a $\neg p$ per assorbimento; il primo è equivalente a $(\neg p \vee \neg q) \wedge (q \vee \neg q)$, quindi a $\neg p \vee \neg q$, e in definitiva

$$(\neg p \vee \neg q) \wedge \neg p,$$

che per l'assorbimento è equivalente a $\neg p$. Questa seconda strada è più lunga, ma è proposta solo per illustrare l'effetto delle varie mosse possibili.

Le forme normali disgiuntive e congiuntive si trovano ai poli estremi di uno spettro su cui si immagini di collocare le proposizioni misurando la loro distanza con il numero di applicazioni delle proprietà distributive necessarie per passare dall'una all'altra. Se si pensasse di decidere se una proposizione in forma normale disgiuntiva è una tautologia applicando il teorema 4.3.3, dovendola prima trasformare in forma congiuntiva, si affronterebbe un compito non inferiore come complessità a quello di costruire la tavola di verità completa (e forse più rischioso, se fatto a mano).

Esercizi

Esercizio 4.3.7. Scrivere la forma normale congiuntiva e disgiuntiva, usando le tavole di verità , delle seguenti proposizioni:

$$(p \vee q \to r) \wedge \neg p \wedge \neg r$$
$$\neg p \to \neg(q \to p)$$
$$(\neg(p \to q) \vee \neg q) \to p$$
$$(q \to r \to p) \to \neg(q \to \neg r \to p)$$

e dire quanti e quali sono i loro modelli.

Esercizio 4.3.8. Per le proposizioni del precedente esercizio, trasformare la forma normale disgiuntiva in quella congiuntiva e viceversa con l'algoritmo FORMA NORMALE.

Scrivere la forma normale disgiuntiva e congiuntiva, usando l'algoritmo FORMA NORMALE, delle seguenti proposizioni:

$$(p \vee q) \to \neg(p \to (q \to r))$$
$$(p \vee q) \to \neg(p \wedge (q \to r))$$
$$p \to (\neg q \vee p \to (r \to p))$$
$$p \oplus (\neg p \oplus q) \to q.$$

Esercizio 4.3.9. Trasformare le leggi logiche della sezione 3.1.1 in forma normale congiuntiva e disgiuntiva.

Esercizio 4.3.10. Osservare che la tavola della proposizione $p \vee r \to \neg p \wedge (q \to r)$ dell'esempio 2.2.2 è uguale a quella di $\neg p$ (se questa è estesa a una tavola a tre entrate p, q, r indipendente da q e r) e trasformare in $\neg p$ la sua forma normale disgiuntiva ottenuta dalla tavola.

Esercizio 4.3.11. Scrivere $\neg p \vee q \to \neg p \wedge q$ in forma normale disgiuntiva e leggerne i modelli. Discutere le relazioni con $p \oplus q$.

Esercizio 4.3.12. Verificare, ai fini dell'applicazione delle trasformazioni con le leggi distributive, che è

$$(A \lor B) \land (C \lor D) \equiv (A \land C) \lor (A \land D) \lor (B \land C) \lor (B \land D)$$

e analogamente

$$(A \land B) \lor (C \land D) \equiv (A \lor C) \land (A \lor D) \land (B \lor C) \land (B \lor D).$$

Esercizio 4.3.13. Verificare che $(\neg p \land q) \lor (p \land q) \equiv q \equiv (\neg p \lor q) \land (p \lor q)$.

Notare che la forma disgiuntiva di sinistra si trasforma nella forma congiuntiva di destra sostituendo $\land$ con $\lor$ e $\lor$ con $\land$.

Trovare un controesempio che mostri che questo non è sempre vero; spiegare perché succede in questo caso.

Esercizio 4.3.14. Verificare come si trasforma, applicando le leggi di De Morgan, la negazione di una forma normale congiuntiva (rispettivamente disgiuntiva) in una forma normale disgiuntiva (rispettivamente congiuntiva).

Esercizio 4.3.15. Spiegare, utilizzando le leggi di De Morgan e la legge della doppia negazione, perché $\mathrm{CNF}(A) \equiv \neg\mathrm{DNF}(\neg A)$ e $\mathrm{DNF}(A) \equiv \neg\mathrm{CNF}(\neg A)$.

L'osservazione fornisce un altro modo per ottenere la forma normale disgiuntiva, o congiuntiva, di una proposizione. Se si vuole ad esempio la forma normale disgiuntiva di A, si può provare a vedere se non sia relativamente facile ottenere $\mathrm{CNF}(\neg A)$; ottenuta questa, la si nega e si applica De Morgan; spesso si evita così l'applicazione ripetuta delle leggi distributive.

Errore frequente: lo studente ha trovato $\mathrm{DNF}(A)$ e per ottenere $\mathrm{CNF}(A)$ nega $\mathrm{DNF}(A)$ e applica De Morgan, ricordando malamente l'esercizio 4.3.15, perché ottiene sì una forma congiuntiva, ma quella della negazione: $\mathrm{CNF}(\neg A)$. È forse il residuo dell'idea di premettere due negazioni, usandone una per trasformare DNF in CNF con De Morgan: $\neg\neg\mathrm{DNF}(A)$, $\neg(\neg\mathrm{DNF}(A))$, $\neg\mathrm{CNF}(\neg A)$. Di quella esterna però ci si dimentica — se si tenesse conto dell'altra negazione, una nuova applicazione di De Morgan riporterebbe a $\mathrm{DNF}(A)$. Due negazioni consecutive non possono creare nulla di nuovo.

Esercizio 4.3.16. In riferimento alle osservazioni del precedente esercizio, trovare la forma normale disgiuntiva e congiuntiva e confrontare i diversi modi per ottenerle, per le proposizioni

$$(p \to q) \to (r \to \neg p)$$
$$p \lor q \to \neg p \lor q$$
$$p \lor (q \land r) \to (\neg r \to p).$$

Esercizio 4.3.17. Discutere e spiegare perché non si adotta $\{\neg, \oplus, \land\}$ per la definizione delle forme normali disgiuntive, nonostante l'esercizio 4.1.6.

Capitolo 5

Calcolo della risoluzione

Il calcolo della risoluzione serve a stabilire l'insoddisfacibilità di insiemi di proposizioni scritte in forma normale congiuntiva.

Il calcolo della risoluzione si esprime nella costruzione di particolari deduzioni, non dissimili in linea di principio da quelle della deduzione naturale, ma che hanno il seguente vantaggio, che esiste una sola regola di inferenza. La ricerca, o costruzione della deduzione, ne viene notevolmente semplificata. Il prezzo che si paga è che un qualsiasi problema semantico, formulato di solito originariamente come $A_1, \ldots, A_n \models B$, richiede un lavoro di *pre-processing* per trasformare il problema e le formule nella forma adatta alla applicazione del calcolo.

5.1 Risoluzione

Ai fini della soddisfacibilità è equivalente considerare una congiunzione o l'insieme dei suoi congiunti, per cui le proposizioni in forma normale congiuntiva si identificano con insiemi di clausole. Un insieme $\mathcal{S}$ di clausole è soddisfacibile se e solo se esiste una i tale che $i \models C$ per ogni $C \in \mathcal{S}$.

Poiché in una clausola c'è solo il connettivo $\vee$ applicato ai letterali, e si possono eliminare ripetizioni, e l'ordine non è rilevante per la commutatività della disgiunzione, una clausola è completamente determinata dall'insieme dei suoi letterali, e spesso è rappresentata come insieme di letterali $\{l_1, \ldots, l_n\}$. Noi manterremo in generale la scrittura di $\vee$ e la rappresentazione come liste, ricordando tuttavia che, in questo paragrafo, una clausola indicherà in generale una qualunque delle clausole equivalenti che si ottengono permutando i letterali. Useremo peraltro in qualche caso, nella trattazione teorica, la notazione insiemistica, comoda per semplicità di espressione; ad esempio la notazione $C \cup \{l\}$ servirà a indicare una clausola che contiene il letterale l, non importa in quale posizione, e $C_1 \cup C_2$ la disgiunzione delle due clausole C_1 e C_2 da cui siano state eliminate le ripetizioni, dovute alla eventuale presenza di uno stesso letterale sia in C_1 sia in C_2. La lettera C in generale indicherà una clausola.

Una clausola unitaria è una clausola con un solo letterale; la terminologia insiemistica suggerisce di considerare per estensione anche la clausola con nessun letterale, o l'insieme vuoto di letterali: questa clausola, che sarebbe $\emptyset$, viene tradizionalmente indicata con $\square$ ed è l'unica clausola che è insoddisfacibile[1]. Infatti $i \models C$ se e solo se esiste un letterale $l \in C$ tale che $i \models l$, e se C è vuota nessuna interpretazione la soddisfa, perché nessuna interpretazione dà il valore 1 a qualche letterale di C[2]. Si noti invece che l'insieme vuoto di clausole è soddisfacibile, perché tutte le sue clausole sono soddisfatte da (una qualunque) i.

Se $\mathcal{S}_1$ è soddisfacibile e $\mathcal{S}_2 \subseteq \mathcal{S}_1$, allora $\mathcal{S}_2$ è soddisfacibile dalle stesse interpretazioni che soddisfano $\mathcal{S}_1$. Per le clausole invece se $C_1 \subseteq C_2$ allora se $i \models C_1$ allora $i \models C_2$.

Conviene indicare con l^c il letterale *complementare* di l, cioè l^c è $\neg p$ se l è p, e l^c è p se l è $\neg p$. Si dice anche che l e l^c formano una coppia complementare.

Definizione 5.1.1 (Regola di risoluzione)**.** Dalle clausole $C_1 \cup \{l\}$ e $C_2 \cup \{l^c\}$ si deriva $C_1 \cup C_2$, schematicamente

$$\frac{C_1 \cup \{l\} \quad C_2 \cup \{l^c\}}{C_1 \cup C_2}$$

Le clausole $C_1 \cup \{l\}$ e $C_2 \cup \{l^c\}$ si chiamano clausole *genitrici* della regola, $C_1 \cup C_2$ la clausola *risolvente*, e l il letterale *risolto*.

Anche l^c si potrebbe chiamare il letterale risolto, ma quando è in gioco una coppia complementare è sufficiente menzionare uno solo dei due. Nelle premesse della regola, C_1 o C_2 o entrambe possono essere vuote.

Con l'iterazione della regola di risoluzione si costruiscono deduzioni, che si possono presentare sia come alberi che come successioni finite di clausole. Come al solito identifichiamo i nodi con le clausole che sono le etichette; gli alberi sono rappresentati ora in modo naturale, con la radice in basso, le foglie in alto; le foglie non hanno predecessori, ogni altro nodo ha due predecessori.

Una *derivazione per risoluzione* della clausola C dall'insieme di clausole $\mathcal{S}$ è un albero etichettato in cui la radice è C, le foglie sono clausole di $\mathcal{S}$ e ogni nodo che non sia una foglia è la risolvente dei due predecessori.

[1]Da non confondere con lo stesso segno $\square$ usato per indicare la fine di una dimostrazione.

[2]Se nel linguaggio c'è $\perp$, la clausola unitaria $\perp$ è insoddisfacibile, e non c'è bisogno di utilizzare $\square$.

Come successione di clausole, una derivazione per risoluzione di una clausola C da $\mathcal{S}$ è una succesione $C_1, \ldots, C_n$ dove C_n è C e ogni C_i o $\in \mathcal{S}$ o è la risolvente di due clausole precedenti della successione.

Se esiste una derivazione di C da $\mathcal{S}$ si dice che C è derivabile da $\mathcal{S}$ e si scrive $\mathcal{S} \vdash C$ (o $\mathcal{S} \vdash_r C$, ma in tutto il capitolo si tratterà solo della risoluzione).

Una derivazione di $\square$ da $\mathcal{S}$ si chiama una *refutazione* di $\mathcal{S}$. La clausola vuota $\square$ finale non può che essere la risolvente di due clausole unitarie l e l^c.

La definizione di derivazione da $\mathcal{S}$ ha senso anche per $\mathcal{S}$ infiniti; ma in ogni derivazione da $\mathcal{S}$ occorrono solo un numero finito di elementi di $\mathcal{S}$.

Esempio 5.1.2. Sia $\mathcal{S} = \{\neg p \lor q, p, \neg q\}$.

$$\begin{tikzcd}
\neg p \lor q \arrow[dr] & & p \arrow[dl] & \\
& q \arrow[dr] & & \neg q \arrow[dl] \\
& & \square &
\end{tikzcd}$$

è una refutazione di $\mathcal{S}$: $\mathcal{S} \vdash \square$.

Esempio 5.1.3. Sia $\mathcal{S} = \{p \lor \neg q \lor r, \neg p \lor \neg q, q, \neg q \lor \neg r\}$.

$$\begin{tikzcd}
p \lor \neg q \lor r \arrow[dr] & & \neg p \lor \neg q \arrow[dl] & & & & & \\
& \neg q \lor r \arrow[dr] & & q \arrow[dl] & & q \arrow[dr] & & \neg q \lor \neg r \arrow[dl] \\
& & r \arrow[drr] & & & & \neg r \arrow[dll] & \\
& & & & \square & & &
\end{tikzcd}$$

è una derivazione di $\square$ da $\mathcal{S}$: $\mathcal{S} \vdash \square$.

Il motivo per cui una derivazione di $\square$ da $\mathcal{S}$ si chiama, opportunamente, una refutazione di $\mathcal{S}$ è che

Teorema 5.1.4 (Correttezza)**.** *Se $\mathcal{S} \vdash \square$, allora $\mathcal{S}$ è insoddisfacibile.*

Dimostrazione. Se i è un'interpretazione che soddisfa $\mathcal{S}$, allora i soddisfa tutte le proposizioni in una derivazione da $\mathcal{S}$, che quindi non può terminare con $\square$.

Infatti se i soddisfa le premesse $C_1 \cup \{l\}$ e $C_2 \cup \{l^c\}$ di una applicazione della regola di risoluzione, allora per la prima genitrice o i soddisfa un letterale di C_1, e quindi $C_1 \cup C_2$, oppure soddisfa l; in questo caso, per la seconda genitrice i deve soddisfare un letterale di C_2, e quindi $C_1 \cup C_2$. $\square$

Prima di dimostrare il viceversa, introduciamo la seguente operazione su insiemi di clausole, con la notazione

$$\mathcal{S}^l = \{C \setminus \{l^c\} : C \in \mathcal{S} \text{ e } l \notin C\}.$$

Analogamente per $\mathcal{S}^{l^c}$, osservando che $(l^c)^c = l$,

$$\mathcal{S}^{l^c} = \{C \setminus \{l\} : C \in \mathcal{S} \text{ e } l^c \notin C\}.$$

L'insieme $\mathcal{S}^l$ si ottiene da $\mathcal{S}$ tenendo le clausole che non contengono né l né l^c, eliminando quelle che contengono l e modificando quelle che contengono l^c con l'eliminazione di l^c.

Lemma 5.1.5. *Se $\mathcal{S}$ è insoddisfacibile, anche $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$ lo sono.*

Dimostrazione. Supponiamo che $\mathcal{S}^l$ sia soddisfatto da i (per $\mathcal{S}^{l^c}$ il discorso è analogo). Poiché né l né l^c occorrono in $\mathcal{S}^l$, i non è definita per l e la si può estendere a i' ponendo $i'(l) = 1$.

Le clausole di $\mathcal{S}$ che sono anche in $\mathcal{S}^l$ perché non contengono né l né l^c e che sono soddisfatte da i continuano a essere soddisfatte da i'. Le clausole di $\mathcal{S}$ che si ottengono da clausole di $\mathcal{S}^l$ reinserendo il letterale l^c che era stato tolto, continuano a essere soddisfatte da i'. Le clausole di $\mathcal{S}$ che non sono in $\mathcal{S}^l$ perché contengono l sono soddisfatte da $i'(l) = 1$. □

Vale anche il viceversa (esercizio).

Se $\mathcal{S}$ è finito, $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$ hanno meno letterali di $\mathcal{S}$. Iterando il passaggio da $\mathcal{S}$ a $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$, determinando $(\mathcal{S}^l)^M$, $(\mathcal{S}^{l^c})^M$ e via riducendo, si perviene a insiemi la cui insoddisfacibilità è facilmente decidibile. L'organizzazione di questo procedimento sui vari insiemi di clausole non è semplice, e useremo $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$ solo per dimostrare, per $\mathcal{S}$ finito, il

Teorema 5.1.6 (Completezza). *Se $\mathcal{S}$ è insoddisfacibile, allora $\mathcal{S} \vdash \square$.*

Dimostrazione. La dimostrazione è per induzione sul numero di letterali degli insiemi di clausole. $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$ hanno meno letterali di $\mathcal{S}$ e sono insoddisfacibili. per ipotesi induttiva, $\mathcal{S}^l \vdash \square$ e $\mathcal{S}^{l^c} \vdash \square$. Se nella refutazione di $\mathcal{S}^l$ si reintroduce l^c nelle clausole di $\mathcal{S}^l$ che sono foglie e che si ottengono da clausole di $\mathcal{S}$ per cancellazione di l^c, questo letterale si trasmette nelle clausole risolventi, e la derivazione diventa una derivazione di l^c da $\mathcal{S}$. Analogamente, se nella refutazione di $\mathcal{S}^{l^c}$ si reintroduce l nelle clausole da cui era stato cancellato, si ottiene una derivazione di l da $\mathcal{S}$.

Accostando queste due derivazioni e aggiungendo una risoluzione tra l e l^c, si ottiene una refutazione di $\mathcal{S}$.

Per la base dell'induzione, si può considerare il caso di zero letterali, e allora $\mathcal{S}$ deve contenere solo la clausola $\square$, che è quindi derivabile da $\mathcal{S}$. □

Prima di provare a costruire una refutazione di $\mathcal{S}$, nell'intento di dimostrarne l'insoddisfacibilità, conviene eseguire su $\mathcal{S}$ alcune semplificazioni, se possibile.

Ad esempio se in qualche clausola di $\mathcal{S}$ occorre un letterale l tale che il suo complementare non occorre in nessuna clausola, allora le clausole contenenti l possono essere eliminate; infatti $\mathcal{S}$ è soddisfacibile se e solo se il sottoinsieme $\mathcal{S}'$ ottenuto in questo modo è soddisfacibile: se $\mathcal{S}$ è soddisfacibile, anche $\mathcal{S}'$ lo è; viceversa, se $\mathcal{S}'$ che non contiene né l né l^c è soddisfatto da i, estendendo i ponendo $i'(l) = 1$ si ha un'interpretazione che soddisfa $\mathcal{S}$.

Se qualche clausola di $\mathcal{S}$ è una tautologia, si può eliminare perché l'insieme ridotto è soddisfacibile se e solo se $\mathcal{S}$ lo è. Ma si può fare di più, chiedendo che in una derivazione da $\mathcal{S}$ ogni volta che l'applicazione della regola di risoluzione produrrebbe una risolvente che è una tautologia, perché in C_1 c'è un letterale l_1 diverso dal letterale risolto, il cui complementare l_1^c occorre in C_2, la risoluzione non venga eseguita. Con questa *restrizione della tautologia*, per cui in una derivazione da $\mathcal{S}$ non deve occorrere in nessun nodo una tautologia, si può dimostrare, rifacendo la dimostrazione del teorema di completezza, che il calcolo della risoluzione resta completo.

Restrizioni del genere sull'applicazione della regola di risoluzione danno origine ai cosiddetti *raffinamenti* della risoluzione.

5.1.1 Risoluzione lineare ordinata

La restrizione della risoluzione lineare ordinata consiste nel chiedere innanzi tutto che ogni risolvente sia immediatamente usata come genitrice di una nuova risoluzione, in cui l'altra genitrice o appartiene all'insieme $\mathcal{S}$ o è una delle clausole ottenute precedentemente nella catena di risoluzioni che ha portato alla presente clausola, in modo che una derivazione da $\mathcal{S}$ assume il seguente aspetto di *derivazione lineare*:

$$\begin{array}{ccc} C_0 & & B_0 \\ \downarrow & \swarrow & \\ C_1 & & B_1 \\ \downarrow & \swarrow & \\ C_2 & & \\ & & \\ \vdots & & \\ \downarrow & \swarrow & \\ C_n & & B_n \\ \downarrow & \swarrow & \\ \square & & \end{array}$$

Le clausole C_i si chiamano *centrali* e tutte salvo la prima, detta *top* e appartenente a $\mathcal{S}$, sono risolventi; le clausole B_i sono dette *laterali* e o appartengono a $\mathcal{S}$ oppure sono una delle C_j con $j < i$. Per queste laterali non è rispettata la condizione che le foglie siano elementi di $\mathcal{S}$, ma la condizione si può ripristinare (distruggendo la linearità) aggiungendo sopra ad esse la derivazione da cui sono dedotte.

Alla struttura lineare della derivazione si aggiunge la condizione che le clausole non sono più considerate come insiemi di letterali, cioè a meno di equivalenze logiche; ogni clausola intesa come disgiunzione è considerata diversa dalle clausole equivalenti che si ottengono per permutazioni dei letterali; le clausole sono successioni ordinate di letterali, inclusa eventualmente la successione vuota, sempre indicata con $\square$.

Si evitano peraltro le ripetizioni di letterali convenendo che quando un letterale occorre più di una volta in una clausola, si cancellando automaticamente tutte le occorrenze salvo la prima (operazione detta anche di *merging left*).

Anche la regola di risoluzione deve essere modificata per adattarsi alla natura ordinata delle clausole, e deve indicare quale letterale scegliere come letterale risolvente e in quale delle due clausole genitrici (che anch'esse sono date in modo ordinato come prima e seconda genitrice). Per comodità di scrittura conveniamo che sia l'ultimo letterale della prima clausola genitrice (ma non cambierebbe nulla con un'altra scelta, ad esempio il primo) nel formulare la

Definizione 5.1.7 (Regola di risoluzione ordinata)**.** Se l è l'ultimo letterale di una prima clausola, ed l^c occorre in una seconda clausola $D_1 \vee l^c \vee D_2$, dove D_1 o D_2 possono essere la clausola vuota, a significare che l^c può essere in una posizione qualunque, allora

$$\frac{C \vee l \quad D_1 \vee l^c \vee D_2}{C \vee D_1 \vee D_2}$$

Si noti che se un letterale di una clausola ha il complementare in un'altra, ma non è all'ultimo posto della sua, non può essere il letterale risolto di una risoluzione delle due clausole: ad esempio $p \vee q$ e $\neg p \vee q$ non possono essere risolte rispetto a p con una risoluzione ordinata.

Una *derivazione per risoluzione lineare ordinata* di C da $\mathcal{S}$ è una derivazione lineare di C da $\mathcal{S}$ in cui le clausole sono considerate ordinate e le risoluzioni sono risoluzioni ordinate tra la centrale e la laterale, nell'ordine.

Esempio 5.1.8. La derivazione nell'Esempio 5.1.2 a p. 109 è una derivazione lineare ma non lineare ordinata. Essa può essere sostituita, per refutare $\mathcal{S} = \{\neg p \vee q, p, \neg q\}$, dalla derivazione lineare ordinata:

$$\begin{tikzcd}[column sep=small]
\neg p \lor q \arrow[dr] & & \neg q \arrow[dl] & \\
& \neg p \arrow[dr] & & p \arrow[dl] \\
& & \square &
\end{tikzcd}$$

Esempio 5.1.9. La refutazione dell'Esempio 5.1.3 di $\mathcal{S} = \{p \lor \neg q \lor r, \neg p \lor \neg q, q, \neg q \lor \neg r\}$ non è lineare. Essa può tuttavia essere sostituita dalla refutazione lineare ordinata:

$$\begin{tikzcd}[column sep=small]
p \lor \neg q \lor r \arrow[dr] & & \neg q \lor \neg r \arrow[dl] & & & \\
& p \lor \neg q \arrow[dr] & & q \arrow[dl] & & \\
& & p \arrow[dr] & & \neg p \lor \neg q \arrow[dl] & \\
& & & \neg q \arrow[dr] & & q \arrow[dl] \\
& & & & \square &
\end{tikzcd}$$

Esempio 5.1.10. Nelle due refutazioni degli esempi precedenti, le clausole laterali sono clausole di $\mathcal{S}$. Nella seguente refutazione lineare ordinata di $\{p \lor q, \neg p \lor q, p \lor \neg q, \neg p \lor \neg q\}$

$$\begin{tikzcd}[column sep=small]
p \lor q \arrow[dr] & & p \lor \neg q \arrow[dl] & & & \\
& p \arrow[dr] & & \neg p \lor q \arrow[dl] & & \\
& & q \arrow[dr] & & \neg p \lor \neg q \arrow[dl] & \\
& & & \neg p \arrow[dr] & & p \arrow[dl] \\
& & & & \square &
\end{tikzcd}$$

l'ultima clausola laterale p è una clausola centrale precedente.

Non sorprenda il secondo esempio, perché vale la completezza per la risoluzione lineare ordinata. La dimostrazione si svolge come quella per la risoluzione ordinaria, manipolando ed attaccando derivazioni che esistono per ipotesi induttiva per insiemi ridotti; questa volta tuttavia le due derivazioni vanno attaccate in serie per formare un albero lineare, e quindi bisogna poter controllare il *top* della seconda. Occorre poter affermare cioè che se un insieme è insoddisfacibile esiste una sua refutazione che ha come *top* una clausola particolare scelta da noi.

Questo è vero, come vedremo, nella forma: se $\mathcal{S}$ è insoddisfacibile e $\mathcal{S} \setminus \{C\}$ è soddisfacibile, allora esiste una refutazione di $\mathcal{S}$ con *top* C. Il risultato è anche operativamente significativo e importante, perché gli insiemi di clausole su cui normalmente si lavora derivano da una situazione del genere: si ha una teoria con assiomi $\{A_1, \ldots, A_n\}$ che si sa o si presume consistente, e si vuole sapere se $A_1, \ldots, A_n \models B$, trasformando il problema in quello della insoddisfacibilità di $\{A_1, \ldots, A_n, \neg B\}$; allora $\neg B$ è il candidato *top* della refutazione (almeno se $A_1, \ldots, A_n, \neg B$ sono clausole).

Per la dimostrazione serve una versione più forte, anche se non operativa, in cui si possa assumere una clausola qualunque come *top*; a questa si arriva con il concetto di *insieme insoddisfacibile minimale*. Un insieme $\mathcal{S}$ di clausole si dice insoddisfacibile minimale se $\mathcal{S}$ è insoddisfacibile e ogni $\mathcal{S}' \subset \mathcal{S}$ è soddisfacibile.

Ogni insieme insoddisfacibile di clausole contiene un sottoinsieme insoddisfacibile minimale (esercizio). Anzi, per ogni insieme $\mathcal{S}$ insoddisfacibile e tale che $\mathcal{S} \setminus \{C\}$ è soddisfacibile, esiste un sottoinsieme insoddisfacibile di $\mathcal{S}$ che contiene C ed è tale che ogni suo sottoinsieme che contiene C è soddisfacibile (esercizio).

Possiamo ora enunciare

Teorema 5.1.11 (Completezza). *Se $\mathcal{S}$ è insoddisfacibile e $\mathcal{S} \setminus \{C\}$ è soddisfacibile, allora esista una refutazione lineare ordinata di $\mathcal{S}$ con top C.*

Dimostrazione. Possiamo supporre che $\mathcal{S}$ sia insoddisfacibile minimale, nel senso detto sopra, vale a dire ogni suo sottoinsieme proprio che contiene C è soddisfacibile.

Sia C della forma $m \vee D \vee l$ (se C è unitaria uguale a l si procede come sotto, identificando M ed l). Si consideri $\mathcal{S}^{m^c}$ che è insoddisfacibile, e si osservi che $\mathcal{S}^{m^c} \setminus \{D \vee l\}$ è soddisfacibile. Infatti se i è un'interpretazione che soddisfa $\mathcal{S} \setminus \{C\}$, ma non C altrimenti soddisferebbe $\mathcal{S}$, allora $i(m) = 0$; allora cancellando m da clausole di $\mathcal{S} \setminus \{C\}$ soddisfatte da i per ottenere $\mathcal{S}^{m^c} \setminus \{D \vee l\}$ si ottengono clausole ancora soddisfatte da i.

Per ipotesi induttiva, esiste una refutazione lineare ordinata di $\mathcal{S}^{m^c}$ con *top* $D \vee l$. Reintroduciamo m nelle clausole da cui era stato tolto, nella stessa posizione in cui era quando è stato tolto, per riavere le stesse clausole di $\mathcal{S}$; ma m è reintrodotto quindi all'inizio del *top*, ed essendo all'inizio m si trasmette lungo l'albero di derivazione restando sempre al primo posto e non interferendo con le risoluzioni compiute. Si ottiene così una derivazione per risoluzione lineare ordinata da $\mathcal{S}$ di m con *top* C.

In $\mathcal{S}$ esiste una clausola che contiene m^c, perché i letterali si presentano sempre in coppie complementari, altrimenti li elimineremmo preliminarmente, e sia $D_1 \vee m^c \vee D_2$. $\mathcal{S} \setminus \{D_1 \vee m^c \vee D_2\}$ contiene C ed è soddisfacibile, ma una j che lo soddisfi deve dare il valore zero a m^c. Se consideriamo ora $\mathcal{S}^m$ abbiamo come sopra che questo è insoddisfacibile e $\mathcal{S}^m \setminus \{D_1 \vee D_2\}$ è soddisfacibile; quindi per ipotesi induttiva esiste una refutazione lineare ordinata di $\mathcal{S}^m$ con *top* $D_1 \vee D_2$. Appendiamo questa alla precedente derivazione di m da $\mathcal{S}$ con il legame

$$
\begin{array}{ll}
\vdots & \\
\downarrow \swarrow & \\
m & D_1 \vee m_c \vee D_2 \\
\downarrow \swarrow & \\
D_1 \vee D_2 & \\
\downarrow & \\
\vdots &
\end{array}
$$

Nella parte sotto $D_1 \vee D_2$ ci sono in generale clausole laterali di $\mathcal{S}^m$; se reintroduciamo m^c otteniamo clausole di $\mathcal{S}$, ma il letterale m^c che passa nelle centrali può a un certo punto essere l'ultimo di una centrale, e bloccare l'applicazione della risoluzione eseguita nella orginaria derivazione con *top* $D_1 \vee D_2$ da $\mathcal{S}^m$; ad esempio potrebbe succedere che da

$$
\begin{array}{ll}
\downarrow & \\
E & F \\
\downarrow \swarrow & \\
G &
\end{array}
$$

si ottiene

$$
\begin{array}{ll}
\downarrow & \\
E & F \vee m^c \\
\downarrow \swarrow & \\
G \vee m^c &
\end{array}
$$

e non si può proseguire con la risoluzione sull'ultimo letterale di G; ma a questo punto possiamo inserire una risoluzione con la clausola centrale precedente m, e ripristinare l'originaria centrale

$$\begin{tikzcd}
{} \arrow[d] & \\
E \arrow[d] & F \vee m^c \arrow[dl] \\
G \vee m^c \arrow[d] & m \arrow[dl] \\
G &
\end{tikzcd}$$

Alla fine si ottiene una derivazione per risoluzione lineare ordinata di $\Box$ da $\mathcal{S}$ con *top* C. $\Box$

Si noti che se C è unitaria $\mathcal{S}^{m^c}$ è $\Box$; in pratica serve solo la seconda derivazione da $\mathcal{S}^m$, con *top* C.

La risoluzione lineare ordinata è compatibile anche con la restrizione della tautologia, perché la dimostrazione si estende a questa ulteriore restrizione: se le refutazioni assunte per ipotesi induttiva soddisfano la restrizione della tautologia, le operazioni eseguite su di esse per combinarle in una refutazione di $\mathcal{S}$ non introducono alcuna tautologia.

Una derivazione lineare può essere ridotta da un albero a un ramo se le frecce che partono dalle clausole laterali sono cancellate e le clausole stesse sono ad esempio scritte a fianco delle frecce che uniscono le due centrali consecutive. Ma allora con un albero si possono rappresentare tutte le derivazioni con *top* fissato, ogni ramo rappresentando una derivazione lineare. Ad esempio, per l'insieme $\{p \vee q \vee r, \neg r \vee q, \neg q, \neg q \vee p, \neg p\}$ tutte le refutazioni lineari ordinate da $\mathcal{S}$ con *top* $p \vee q \vee r$ sono rappresentate dall'albero

$$\begin{tikzcd}
& p \vee q \vee r \arrow[d] & \\
& p \vee q \arrow[dl, "\neg q"'] \arrow[dr, "\neg q \vee p"] & \\
p \arrow[d, "\neg p"'] & & p \arrow[d, "\neg q \vee p"] \\
\Box & & \Box
\end{tikzcd}$$

Senza il vincolo dell'ordine in generale l'albero delle derivazioni sarebbe molto più folto.

Esercizio 5.1.12. Descrivere l'insieme $\mathcal{S}$ di clausole associate alla forma normale congiuntiva delle seguenti proposizioni:

$$(p \to q \lor r) \land \neg r \land (q \to \neg p)$$
$$\neg(p \to q) \lor (p \land \neg q)$$
$$(p \to q) \to ((\neg p \to q) \to \neg q).$$

Esercizio 5.1.13. Determinare $\mathcal{S}^p$ e $\mathcal{S}^{\neg p}$ per i seguenti insiemi di clausole:

$$\{p \lor q, \neg p \lor q, q \lor r, \neg p \lor r\}$$
$$\{p \lor \neg q \lor \neg r, \neg p \lor q, \neg p \lor r, p \lor \neg q\}$$
$$\{\neg p \lor q, \neg q \lor r, p \lor \neg r\}.$$

Esercizio 5.1.14. Dimostrare che se $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$ sono insoddisfacibili, anche $\mathcal{S}$ lo è.

Esercizio 5.1.15. Spiegare come devono essere fatti gli insiemi di clausole a cui si perviene iterando le riduzioni $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$ perché $\mathcal{S}$ sia insoddisfacibile.

Esercizio 5.1.16. Verificare l'insoddisfacibilità o meno dei seguenti insiemi di clausole calcolando iterativamente le riduzioni $\mathcal{S}^l$ e $\mathcal{S}^{l^c}$:

$$\{p \lor r, q \lor \neg r, \neg q, \neg p \lor t, \neg u, u \lor \neg t\}$$
$$\{p \lor \neg q \lor r, q \lor r, \neg p \lor r, q \lor \neg r, \neg q\}$$
$$\{p \lor q \lor r, \neg p \lor q, \neg q\}.$$

Esercizio 5.1.17. Cercare una refutazione per risoluzione lineare dei seguenti insiemi di clausole:

$$\{p \lor r, q \lor \neg r, \neg q, \neg p \lor t, \neg u, u \lor \neg t\}$$
$$\{p \lor \neg q \lor r, q \lor r, \neg p \lor r, q \lor \neg r, \neg q\}$$

Esercizio 5.1.18. Trovare una refutazione lineare ordinata dei seguenti insiemi di clausole:

$$\{p \lor r, q \lor \neg r, \neg q, \neg p \lor t, \neg u, u \lor \neg t\}$$
$$\{p \lor \neg q \lor r, q \lor r, \neg p \lor r, q \lor \neg r, \neg q\}$$
$$\{p \lor q \lor r, \neg p \lor q, \neg q, \neg r \lor q\}.$$

Esercizio 5.1.19. Cercare una refutazione per risoluzione non lineare dei seguenti insiemi di clausole:

$$\{p \lor r, q \lor \neg r, \neg q, \neg p \lor t, \neg u, u \lor \neg t\}$$
$$\{p \lor \neg q \lor r, q \lor r, \neg p \lor r, q \lor \neg r, \neg q\}.$$

Esercizio 5.1.20. Verificare che per l'insieme insoddisfacibile $\{p \vee q \vee r, \neg p \vee q, \neg q, \neg r \vee q, \neg p \vee \neg r\}$ non c'è una refutazione lineare ordinata con *top* $\neg p \vee \neg r$ che rispetti anche la restrizione della tautologia, e dare una spiegazione.

Esercizio 5.1.21. Trovare tutte le refutazioni lineari ordinate di $\mathcal{S} = \{p \vee \neg q \vee r, \neg p \vee \neg q, q, \neg q \vee \neg r\}$.

Esercizio 5.1.22. Rappresentare per ogni clausola di $\mathcal{S}$ dell'esercizio precedente l'albero delle derivazioni lineari, anche non ordinate, da $\mathcal{S}$ che hanno quella clausola come *top*.

Esercizio 5.1.23. Dimostrare che ogni insieme insoddisfacibile di clausole contiene un sottoinsieme insoddisfacibile minimale.

5.2 Clausole di Horn e programmazione logica

Una clausola si chiama *clausola di Horn* se contiene al più un letterale positivo. Le clausole si possono classificare in *negative* se tutti i loro letterali sono negativi, *positive* se tutti i loro letterali sono positivi, e altrimenti *miste*.

Una clausola di Horn può dunque essere negativa, e allora si chiama anche *goal*, per il motivo che vedremo, oppure positiva, ma allora è unitaria, e si chiama anche *fatto*, oppure mista, ma con un solo letterale positivo, e si chiama anche *legge*.

Un insieme di clausole di Horn, se è insoddisfacibile, non può contenere soltanto leggi e fatti, perché altrimenti l'interpretazione *positiva*, quella che dà il valore 1 a tutte le lettere, lo soddisferebbe; deve contenere quindi almeno un *goal*. Analogamente non può contenere solo clausole negative o miste, altrimenti l'interpretazione *negativa*, quella che dà il valore 0 a tutte le lettere, lo soddisferebbe.

Per gli insiemi di clausole di Horn è completa una ulteriore restrizione della risoluzione lineare ordinata, quella della *risoluzione a input*, o input-risoluzione (che si potrebbe definire anche solo sulla base della risoluzione lineare), data dalla

Definizione 5.2.1. Una derivazione per risoluzione lineare ordinata si dice una input-derivazione se nessuna clausola laterale è uguale a una centrale precedente.

Una input-derivazione di $\square$ da $\mathcal{S}$ si chiama una input-refutazione di $\mathcal{S}$.

Il vantaggio di una input-derivazione su una derivazione lineare ordinata è che non c'è bisogno di registrare e conservare in memoria, in un'esecuzione meccanica, le varie clausole centrali generate, ma solo l'ultima, insieme alle clausole originali di $\mathcal{S}$.

Teorema 5.2.2 (Completezza)**.** *Se $\mathcal{S}$ è un insieme insoddisfacibile di clausole di Horn, esiste una input-refutazione di $\mathcal{S}$.*

Dimostrazione. Si consideri una refutazione lineare ordinata di $\mathcal{S}$ che abbia come *top* un *goal.* una tale derivazione esiste, per la completezza della risoluzione lineare ordinata, e supponendo $\mathcal{S}$ insoddisfacibile minimale, o di essersi ristretti a un suo sottonsieme insoddisfacibile minimale, esiste con *top* qualunque; ma ogni insieme insoddisfacibile di clausole di Horn contiene almeno un *goal.* Scegliamo un *top* negativo.

La clausola *top* negativa non può che essere risolta con una clausola mista o positiva, ma la risolvente è di nuovo negativa. Iterando, tutte le centrali sono negative, e nessuna può essere risolta con una centrale precedente, anch'essa negativa.

Quindi una refutazione lineare ordinata di $\mathcal{S}$ con *top* una clausola negativa soddisfa automaticamente la restrizione dell'input. $\square$

Si potrebbe definire una nozione duale rispetto alle clausole di Horn, clausole in cui occorre al più un letterale negativo. Anche per queste vale la completezza della input-risoluzione, con derivazioni che hanno come *top* una clausola positiva.

Nella programmazione logica si usano clausole di Horn, di linguaggi predicativi, ma con un calcolo soggiacente che può essere illustrato con la logica proposizionale, e non è altro che la input-risoluzione, in altra notazione.

Si torna ad usare il condizionale, anche se rovesciato, e una clausola mista $\neg q_n \vee \ldots \vee \neg q_1 \vee p$, che è logicamente equivalente a $q_n \wedge \ldots \wedge q_1 \to p$, sarà rappresentata da

$$p \leftarrow q_1, \ldots, q_n$$

anche se p fosse in un posto qualunque. p si chiama *testa* e $q_1, \ldots, q_n$ *corpo* della legge.

Quindi per uniformità di notazione una clausola unitaria p sarà rappresentata da

$$p \leftarrow$$

e un *goal* $\neg q_n \vee \ldots \vee \neg q_1$ da

$$\leftarrow q_1, \ldots, q_n$$

tralasciamo altri dettagli zuccorosi dei veri e propri linguaggi di programmazione come PROLOG, come un punto alla fine di ogni clausola o un punto interrogativo per il *goal*. Un insieme $\mathcal{P}$ di clausole di Horn positive o miste scritte nel modo indicato si chiama anche *programma*.

Esempio 5.2.3. L'insieme di clausole di Horn $\mathcal{S} = \{p \vee \neg r \vee \neg q, q, r\}$ si rappresenta con il programma

$$\begin{array}{l} q \leftarrow \\ r \leftarrow \\ p \leftarrow q, r \end{array}$$

Un'*interrogazione* di un programma si realizza chiedendo se una proposizione atomica, o una congiunzione di proposizioni atomiche, sono conseguenza logica del programma. questo equivale a chiedere se aggiungendo la loro negazione all'insieme di clausole del programma si ottiene un insieme insoddisfacibile. Al programma si aggiunge pertanto l'interrogazione nella forma di un *goal* semplice, la negazione di una lettera, ovvero una clausola negativa.

Esempio 5.2.4. Il precedente programma può essere interrogato aggiungendo il *goal* $\neg p$, ottenendo

$$\begin{array}{l} q \leftarrow \\ r \leftarrow \\ p \leftarrow q, r \\ \leftarrow p \end{array}$$

Il teorema sulla completezza della input-risoluzione suggerisce di partire per una input-refutazione dal *goal*. Nella notazione della programmazione logica, che seguiamo sull'esempio di sopra, la risoluzione di $\neg p$ con la clausola $p \vee \neg r \vee \neg q$, che dà origine alla nuova clausola negativa, nuovo *goal*, $\neg r \vee \neg q$, viene sostituita dalla identificazione del p del *goal* $\leftarrow p$ con la testa di una legge o con un fatto, in questo caso la legge $p \leftarrow q, r$, e nello spostamento del problema del *soddisfacimento* del *goal* p con quello del soddisfacimento del nuovo *goal* multiplo $\leftarrow q, r$. Formalmente, p viene rimpiazzato in $\leftarrow p$ dal corpo q, r della clausola con la cui testa p si identifica. Si dice anche, in vista dell'estensione ai linguaggi predicativi, che p *unifica* con la testa di una clausola.

Un *goal* multiplo si soddisfa cercando di soddisfare il primo *sottogoal* q con lo stesso procedimento di prima, che corrisponde a cercare di risolvere la clausola $\neg r \vee \neg q$ rispetto al suo ultimo letterale. Poiché q unifica con il fatto $q \leftarrow$, q è soddisfatto, e il *goal* $\leftarrow q, r$ si riduce, sostituendo a q il corpo del fatto $q \leftarrow$, che è vuoto, a $\leftarrow r$. Anche questo *sottogoal* r unifica con il fatto

$r \leftarrow$, ed è perciò soddisfatto. Il *goal* diventa $\leftarrow$, o la clausola vuota, e il *goal* originario è soddisfatto.

Se il procedimento non arrivasse a questa conclusione, ma si interrompesse con qualche *sottogoal* non soddisfatto, o entrasse in ciclo, non si direbbe ancora che il *goal* $\leftarrow p$ non è soddisfatto, o *fallisce*; si esplorerebbero prima altre possibilità di unificazione dei vari *sottogoal*, e solo se tutte fallissero si direbbe che il *goal* è fallito, il che corrisponde alla non esistenza di una input-refutazione con il *goal* come *top*, o alla soddisfacibilità delle clausole del programma più $\neg p$.

Si faccia attenzione che dire nel contesto della programmazione che un *goal* $\leftarrow p$ è soddisfatto dal programma $\mathcal{P}$ significa dire nel contesto logico che l'insieme $\mathcal{P} \cup \{\neg p\}$ è insoddisfacibile

Ad esempio, per il programma

$$\begin{aligned} p &\leftarrow q \\ q &\leftarrow u \\ q &\leftarrow r \\ r &\leftarrow \end{aligned}$$

il *goal* $\leftarrow p$ unifica con la prima clausola e il nuovo *goal* è $\leftarrow q$, che unifica con la seconda, ma diventa $\leftarrow u$ e la ricerca si interrompe. Se si torna indietro e si unifica $\leftarrow q$ con la testa della terza clausola, si ottiene $\leftarrow r$ che è soddisfatto unificando con $r \leftarrow$.

Le due strade seguite corrispondono ai due rami del seguente albero delle derivazioni lineari ordinate da $\{p \vee \neg q, q \vee \neg u, q \vee \neg r, r, \neg p\}$ con *top* $\neg p$:

$$\begin{tikzcd}[column sep=small]
 & \neg p \arrow[d, "p \vee \neg q"] & \\
 & \neg q \arrow[dl, "q \vee \neg u"'] \arrow[dr, "q \vee \neg r"] & \\
\neg u & & \neg r \arrow[d, "r"] \\
 & & \square
\end{tikzcd}$$

Invece il programma

$$\begin{aligned} p &\leftarrow q, r \\ q &\leftarrow r \\ q &\leftarrow p \end{aligned}$$

interrogato con il *goal* $\leftarrow p$ esplora due strade, una che fallisce con il *sottogoal* $\leftarrow r$, l'altra che entra in ciclo, in corrispondenza al seguente albero delle derivazioni da $\{p \vee \neg r \vee \neg q, q \vee \neg r, q \vee \neg p\}$ con *top* $\neg p$,

$$\begin{tikzcd}
& \neg p \arrow[d, "p \vee \neg r \vee \neg q"] & \\
& \neg r \vee \neg q \arrow[dl, "q \vee \neg r"'] \arrow[dr, "q \vee \neg p"] & \\
\neg r & & \neg r \vee \neg p \arrow[d] \\
& & {}
\end{tikzcd}$$

dove il ramo di sinistra è terminato e quello di destra con $p \vee \neg r \vee \neg q$ dà origine a un ciclo infinito.

Esempio 5.2.5. L'insieme di clausole di Horn $\mathcal{S} = \{p \vee \neg r \vee \neg q, q \vee \neg t \vee \neg u, r \vee \neg u, u, t, q\}$ è soddisfacibile, con l'intepretazione positiva, in quanto non ci sono clausole negative. $\mathcal{S} \cup \{\neg p\}$ è insoddisfacibile. Lo verifichiamo nella forma della programmazione logica considerando il programma associato a $\mathcal{S}$:

$$\begin{array}{l}
p \leftarrow q, r \\
q \leftarrow u, t \\
r \leftarrow u \\
u \leftarrow \\
t \leftarrow \\
q \leftarrow
\end{array}$$

a cui aggiungiamo il *goal*

$$\leftarrow p$$

e calcoliamo la successione dei nuovi *goal*

$$\begin{array}{rll}
1 & p \leftarrow q, r & \\
2 & q \leftarrow u, t & \\
3 & r \leftarrow u & \\
4 & u \leftarrow & \\
5 & t \leftarrow & \\
6 & q \leftarrow & \\
7 & \leftarrow p & p \text{ unifica con } 1 \\
8 & \leftarrow q, r & q \text{ unifica con } 2 \\
9 & \leftarrow u, t, r & u \text{ unifica con } 4 \\
10 & \leftarrow t, r & t \text{ unifica con } 5 \\
11 & \leftarrow r & r \text{ unifica con } 3 \\
12 & \leftarrow u & u \text{ unifica con } 4 \\
13 & \leftarrow &
\end{array}$$

Si noti che al passo 8 q unifica anche con 6, per cui si sarebbe potuto continuare con

$$\begin{array}{rl} 9 & \leftarrow r \\ 10 & \leftarrow u \\ 11 & \leftarrow \end{array}$$

Queste sono le due input-refutazioni dell'insieme

$$\mathcal{S} = \{p \vee \neg r \vee \neg q,\ q \vee \neg t \vee \neg u, r \vee \neg u, u, t, q, \neg p\}$$

con *top* $\neg p$. Il motore inferenziale dei linguaggi di programmazione logica si basa tuttavia oltre che sulla regola di risoluzione anche su alcune restrizioni pratiche, per ragioni di controllo; ad esempio le clausole del programma sono date in modo ordinato, e per soddisfare un *sottogoal* le clausole di programma sono passate in rassegna nell'ordine fissato, a cercare una clausola con testa che unifichi con il *sottogoal*, e viene scelta la prima che si incontra. Questo vincolo distrugge la completezza, perchè ad esempio il programma

$$\begin{array}{l} p \leftarrow q \\ q \leftarrow p \\ p \leftarrow \end{array}$$

entra in ciclo con il *goal* $\leftarrow p$ e il *goal* non è soddisfatto, nonostante p sia conseguenza logica di $\{p \to q, q \to p, p\}$. Per via di queste e altre restrizioni, si impone un'attenta scrittura ordinata dei programmi, o l'introduzione di comandi permessi dal linguaggio che impediscano di tornare a clausole già considerate.

Il procedimento inferenziale sopra descritto, nella notazione della programmazione logica, corrispondente alla input-risoluzione, si chiama anche nella letteratura specifica SLD-*risoluzione*, SLD per *Selection Left for Definite clauses*: "clausole definite" è un altro nome per le clausole di Horn; la "scelta a sinistra" si riferisce alla scelta di prendere in considerazione sempre il primo, a sinistra, *sottogoal* di un *goal* multiplo; la scelta corrisponde a quella della risoluzione ordinata di considerare l'ultimo, a destra, letterale come letterale da risolvere, secondo la rappresentazione del *goal* $\neg q_n \vee \ldots \vee \neg q_1$ come $\leftarrow q_1, \ldots, q_n$.

Esercizio 5.2.6. Verificare e spiegare perché non può esistere alcuna input-refutazione dell'insieme $\mathcal{S} = \{p \vee q, \neg p \vee q, \neg q \vee p, \neg p \vee \neg q\}$ dell'esempio di p. 116.

Esercizio 5.2.7. Verificare che l'insieme insoddisfacibile di clausole $\mathcal{S} = \{p \vee \neg r, q \vee \neg r, r \vee \neg q \vee p, \neg p \vee \neg q, q \vee r\}$ non ammette una input-refutazione.

Esercizio 5.2.8. Spiegare perché l'insieme $\{p \vee q, \neg p \vee \neg q, p, q\}$ pur non essendo un insieme di clausole di Horn ammette una input-refutazione.

Esercizio 5.2.9. Trovare una input-refutazione dell'insieme $\mathcal{S} = \{\neg p \vee \neg q \vee \neg r, \neg p \vee \neg q \vee r, q \vee \neg u, u \vee \neg p, p\}$.

Trovare una input-refutazione dell'insieme $\mathcal{S} = \{p \vee q \vee r, \neg r \vee p \vee q, \neg q \vee p, \neg p\}$ con *top* $p \vee q \vee r$ e spiegare perché è prevedibile che esista.

Esercizio 5.2.10. Trovare tutte le input-refutazioni dell'insieme $\mathcal{S} = \{\neg p \vee \neg q \vee \neg r, \neg p \vee r \vee \neg q, q \vee \neg r, r \vee \neg p, \neg q \vee \neg r, p\}$ e rappresentarle con gli alberi di tutte le input-derivazioni con *top* fissato.

Esercizio 5.2.11. Riscrivere gli insiemi di clausole di Horn dei precedenti esercizi nella notazione della programmazione logica e svolgerli nella forma di interrogazioni di programmi.

Esercizio 5.2.12. Dato il programma

$$\begin{array}{l} p \leftarrow q, t \\ u \leftarrow r \\ r \leftarrow p, q \\ t \leftarrow q \\ q \leftarrow \end{array}$$

dimostrare che il *goal* $\leftarrow p$ è soddisfatto, e che anche $\leftarrow q, r$ lo è.

Esercizio 5.2.13. Dimostrare che anche i *goal* $\leftarrow u$ e $\leftarrow t$ sono soddisfatti, e che questo significa che non solo l'interpretazione positiva soddisfa il programma, ma che è l'unica che lo soddisfa.

Esercizio 5.2.14. Dimostrare che se in un programma ci sono solo fatti, solo i *goal* che unificano con quei fatti sono soddisfatti, e che se in un programma ci sono solo leggi nessun *goal* è soddisfatto; questo significa che per ogni *goal* $\leftarrow p$ l'insieme di $\neg p$ e delle clausole del programma è soddisfacibile; quale è l'interpretazione che lo soddisfa?

Esercizio 5.2.15. Verificare che per il programma

$$\begin{array}{l} u \leftarrow r \\ r \leftarrow p, u \\ p \leftarrow u \\ u \leftarrow \end{array}$$

il *goal* $\leftarrow u$ non è soddisfatto se si impone di prendere in considerazione le clausole di programma solo nell'ordine in cui sono scritte.

Capitolo 6

Alberi di refutazione

6.1 Il metodo

Un altro metodo per rispondere alle domande sulla verità logica o sulla insoddisfacibilità delle proposizioni, e che risulta più efficiente (almeno sui casi di dimensione ridotta) della ricerca esaustiva offerta dalla costruzione delle tavole di verità, è il metodo degli *alberi di refutazione*[1].

Il nome deriva dal fatto che il procedimento è impostato, per rispondere alla domanda sulla verità logica, secondo la *ricerca del controesempio*: si cerca di scoprire se esiste un'interpretazione che falsifichi la proposizione. Il metodo ha la proprietà che o la trova, se esiste, e quindi fornisce un'interpretazione in cui la negazione della proposizione è vera (controesempio: la proposizione è falsa) oppure mostra che non è possibile che esista, e quindi la proposizione è una tautologia.

Più in generale, il metodo serve a stabilire se esista o no un'interpretazione che soddisfa una proposizione composta, non partendo dal basso dalle possibili interpretazioni delle lettere (*bottom up*) ma dall'alto, dalla proposizione data, scendendo verso le sottoproposizioni componenti (*top down*); nel processo, si accumulano condizioni necessarie che l'ipotetica interpretazione, se esiste e soddisfa la radice, dovrebbe pure soddisfare — nel senso di quali altre proposizioni essa dovrebbe soddisfare o no — fino alle condizioni necessarie riguardanti le proposizioni atomiche; queste, se non sono incompatibili tra di loro, si traducono in condizioni sufficienti per la definizione dell'interpretazione.

Gli alberi di refutazione possono dunque essere usati anche per rispondere alle altre domande semantiche, ad esempio quella sulla soddisfacibilità.

[1] Altri nomi usati, insieme a qualche variante di presentazione, sono quelli di *alberi semantici*, oppure di *tableaux* semantici.

Gli alberi di refutazione sono alberi etichettati con proposizioni. Identifichiamo per comodità di scrittura i nodi con le loro etichette. Nella radice è una proposizione, di cui si vuole sapere se esiste un modello. L'albero è sviluppato secondo la seguente procedura.

Ad ogni stadio, si saranno già prese in considerazione alcune proposizioni, messe tra parentesi quadre o segnate con un asterisco, e ne resteranno da considerare altre. Se sono già state considerate tutte, l'albero è *terminato*; se no, si prende in esame una proposizione A non ancora considerata, e a seconda della sua forma si prolunga l'albero nel modo seguente, dopo aver segnato A e aver notato quali sono i rami non chiusi che passano per A, dove un ramo si dice *chiuso* se su di esso occorre sia una proposizione sia la sua negazione:

- Se A è una proposizione senza connettivi, non si fa nulla (si va al passo successivo).

- Se A è $B \wedge C$, alla fine di ogni ramo non chiuso passante per A si appendono alla foglia due nodi in serie etichettati con B e C, come nello schema:

$$\begin{array}{c} [B \wedge C] \\ \vdots \\ \downarrow \\ F \\ \downarrow \\ B \\ \downarrow \\ C \end{array}$$

- Se A è $B \vee C$, alla fine di ogni ramo non chiuso passante per A si aggiunge alla foglia una diramazione con due nodi B e C, come nello schema:

$$\begin{array}{ccc} & [B \vee C] & \\ & \vdots & \\ & \downarrow & \\ & F & \\ \swarrow & & \searrow \\ B & & C \end{array}$$

  con l'ovvia generalizzazione (qui e nella prececente regola) che se si tratta di congiunzioni o disgiunzioni generalizzate si appendono in serie o rispettivamente si fanno diramazioni con tanti nodi quante sono le sottoproposizioni immediate.

- Se A è $B \to C$, alla fine di ogni ramo non chiuso passante per A si aggiunge alla foglia una diramazione con due nodi $\neg B$ e C, come nello schema:

$$\begin{array}{ccc} & [B \to C] & \\ & \vdots & \\ & \downarrow & \\ & F & \\ \swarrow & & \searrow \\ \neg B & & C \end{array}$$

- Se A è $\neg B$ e B non ha connettivi, non si fa nulla.

- Se A è della forma $\neg B$ e B è $\neg C$, al fondo di ogni ramo non chiuso passante per A si appende alla foglia il successore C, come nello schema:

$$\begin{array}{c} [\neg\neg C] \\ \vdots \\ \downarrow \\ F \\ \downarrow \\ C \end{array}$$

- Se A è della forma $\neg B$ e B è $B_1 \lor B_2$, alla fine di ogni ramo non chiuso passante per A si aggiungono alla foglia due nodi in serie $\neg B_1$ e $\neg B_2$, come nello schema:

$$\begin{array}{c} [\neg(B_1 \lor B_2)] \\ \vdots \\ \downarrow \\ F \\ \downarrow \\ \neg B_1 \\ \downarrow \\ \neg B_2 \end{array}$$

  con l'ovvia generalizzazione se B è una disgiunzione generalizzata.

- Se A è della forma $\neg B$ e B è $B_1 \to B_2$, alla fine di ogni ramo non chiuso passante per A si appendono alla foglia due successori in serie

B_1 e $\neg B_2$, come nello schema:

$$\begin{array}{c} [\neg(B_1 \to B_2)] \\ \vdots \\ \downarrow \\ F \\ \downarrow \\ B_1 \\ \downarrow \\ \neg B_2 \end{array}$$

- Se A è della forma $\neg B$ e B è $B_1 \wedge B_2$, alla fine di ogni ramo non chiuso passante per A si aggiunge alla foglia una diramazione con due nodi $\neg B_1$ e $\neg B_2$, come nello schema:

$$\begin{array}{ccc} & [\neg(B_1 \wedge B_2)] & \\ & \vdots & \\ & \downarrow & \\ & F & \\ \swarrow & & \searrow \\ \neg B_1 & & \neg B_2 \end{array}$$

Ovviamente se per il nodo in considerazione non passa alcun ramo non chiuso, non si fa nulla. Dalla formulazione è chiaro che quando tutti i rami sono chiusi il procedimento termina, anche se non tutte le proposizioni sono state considerate, e in tal caso l'albero si considera terminato e si dice *chiuso*.

Non diamo le regole per il bicondizionale (esercizio) perché non sarebbero altro che l'adattamento di quelle che derivano dal fatto che $p \leftrightarrow q$ è equivalente a $(p \to q) \wedge (q \to p)$. Lo stesso per $\oplus$, ma si preferisce eliminare prima questi connettivi (comunque, si diano le regole per $\oplus$ per esercizio), e questa è l'unica preparazione o trasformazione che si fa sulle proposizioni; altrimenti si prendono così come sono, e questo è un vantaggio del metodo, nessun *pre-processing*.

Si leggano con attenzione le regole, cogliendone tutte le informazioni e i vincoli: ad esempio, quando si lavora su di un nodo, si aggiungono proposizioni su *tutti* i rami *passanti* per quel nodo, ma *non* sugli altri.

Esempio

1. Consideriamo la proposizione $\neg((\neg p \vee q) \wedge p \to q)$ che mettiamo nella radice dell'albero

$$\neg((\neg p \vee q) \wedge p \to q)$$

2. Lavorando su di esso, che è la negazione di un condizionale, otteniamo

$$\begin{array}{c}
[\neg((\neg p \lor q) \land p \to q)] \\
\downarrow \\
(\neg p \lor q) \land p \\
\downarrow \\
\neg q
\end{array}$$

3. Lavorando su $(\neg p \lor q) \land p$ otteniamo

$$\begin{array}{c}
[\neg((\neg p \lor q) \land p \to q)] \\
\downarrow \\
[(\neg p \lor q) \land p] \\
\downarrow \\
\neg q \\
\downarrow \\
\neg p \lor q \\
\downarrow \\
p
\end{array}$$

4. Lavorando prima su $\neg q$, senza alcun effetto, e poi su $\neg p \lor q$

$$\begin{array}{c}
[\neg((\neg p \lor q) \land p \to q)] \\
\downarrow \\
[(\neg p \lor q) \land p] \\
\downarrow \\
[\neg q] \\
\downarrow \\
[\neg p \lor q] \\
\downarrow \\
p \\
\swarrow \quad \searrow \\
\neg p \qquad\quad q\,.
\end{array}$$

Non è neanche necessario indicare che si sono presi in considerazione le restanti proposizioni, perché il loro effetto è nullo. L'albero è chiuso, perché su uno dei sue due rami occorrono p e $\neg p$, e sull'altro occorrono q e $\neg q$.

Se si deve interpretare come è stato ottenuto un albero sviluppato, è di aiuto che sia segnato a fianco di ogni proposizione l'ordine in cui è stata presa in considerazione, come in

$$
\begin{array}{c}
[\neg((\neg p \lor q) \land p \to q)]_1 \\
\downarrow \\
[(\neg p \lor q) \land p]_2 \\
\downarrow \\
[\neg q]_3 \\
\downarrow \\
[\neg p \lor q]_4 \\
\downarrow \\
p \\
\swarrow \quad \searrow \\
\neg p \qquad\quad q\,.
\end{array}
$$

Esempio

$$
\begin{array}{c}
\neg((p \land q) \lor (\neg p \land q) \lor \neg q)_1 \\
\downarrow \\
\neg(p \land q)_3 \\
\downarrow \\
\neg(\neg p \land q)_4 \\
\downarrow \\
\neg\neg q_2 \\
\downarrow \\
q \\
\swarrow \quad \searrow \\
\begin{array}{cc}
\begin{array}{c}
\neg p \\
\swarrow \quad \searrow \\
\begin{array}{cc}
\begin{array}{c} \neg\neg p \\ \text{chiuso} \end{array} & \neg q
\end{array}
\end{array}
&
\begin{array}{c} \neg q \\ \text{chiuso} \\ \\ \end{array}
\end{array}
\end{array}
$$

dove il ramo di destra con foglia $\neg q$ non è sviluppato con

$$
\begin{array}{c}
\neg q \\
\swarrow \quad \searrow \\
\neg\neg p \qquad \neg q
\end{array}
$$

come dovrebbe essere per il lavoro su $\neg(\neg p \land q)$, perché il ramo è già chiuso; il ramo di sinistra non è prolungato con

$$
\begin{array}{c}
\neg\neg p \\
\downarrow \\
p
\end{array}
$$

perché anch'esso chiuso.

Avvertenza 6.1.1. Non si confondano gli alberi di refutazione con gli alberi sintattici. L'albero sintattico di una proposizione contiene solo le sottoproposizioni di quella data, l'albero di refutazione anche altre. Ad esempio

$$\begin{array}{ccccccc}
 & & & p \lor q \to p \land q & & & \\
 & & \swarrow & & \searrow & & \\
 & p \lor q & & & & p \land q & \\
 \swarrow & & \searrow & & \swarrow & & \searrow \\
 p & & q & & p & & q
\end{array}$$

è l'albero sintattico di $p \lor q \to p \land q$. Il suo albero di refutazione è invece

$$\begin{array}{ccccc}
 & & p \lor q \to p \land q & & \\
 & \swarrow & & \searrow & \\
 \neg(p \lor q) & & & & p \land q \\
 \downarrow & & & & \downarrow \\
 \neg p & & & & p \\
 \downarrow & & & & \downarrow \\
 \neg q & & & & q
\end{array}$$

dove, oltre a una struttura diversa, compaiono proposizioni come $\neg p$, $\neg q$, $\neg(p \lor q)$.

In alcune presentazioni gli alberi di refutazione sono alberi i cui nodi sono proposizioni etichettate con i simboli V ed F, per "vero" e "falso". La corrispondenza con la presente versione si ottiene sostituendo $V\,A$ con A e $F\,A$ con $\neg A$, e viceversa.

Ad esempio le due regole per la congiunzione diventano:

($V\land$) Quando si lavora su $V(A \land B)$ (al fondo di ogni ramo ...) si appendono in serie $V\,A$ e $V\,B$.

($F\land$) Quando si lavora su A $F(A \land B)$ (al fondo di ogni ramo ...) si appende una diramazione con con $F\,A$ e $F\,B$.

6.1.1 Correttezza e completezza

Il primo problema con ogni algoritmo è quello della terminazione, in particolare per gli algoritmi di decisione; se l'algoritmo non si ferma sempre, con una risposta, dopo un numero finito di passi, non ci si può affidare ad esso per decidere le questioni che interessano (nel senso di lanciarlo e stare ad aspettare).

Lemma 6.1.2 (Terminazione)**.** *La costruzione dell'albero di refutazione inizializzato con una proposizione termina sempre in un numero finito di passi.*

Dimostrazione. Se ad ogni stadio si lavora su una proposizione di quelle che hanno altezza massima n tra quelle non ancora considerate, l'applicazione delle regole fa sì che dopo un numero finito di passi tutte quelle di altezza n siano state considerate, e l'altezza massima delle proposizioni non ancora considerate sia quindi $< n$. Infatti le proposizioni introdotte nell'albero con le regole hanno tutte altezza minore della proposizione che governa la regola, salvo il caso di $B \to C$, per cui si introducono $\neg B$ e C, e $\neg B$ può avere la stessa altezza di $B \to C$ (quando? esercizio); ma la successiva applicazione di una delle regole per proposizioni negate a $\neg B$, che si può eseguire subito, la sostituisce con proposizioni di altezza minore. □

Anche se dunque nel corso del procedimento il numero di proposizioni nei nodi dell'albero cresce con il crescere dell'albero, diminuisce quello delle proposizioni di altezza massima, e dopo un numero finito di passi ci saranno solo proposizioni di altezza minima, senza connettivi, non ancora considerate, e a quel punto il processo termina, se non è terminato prima per la chiusura dell'albero.

Per quel che riguarda la correttezza e la completezza del metodo, qualche perplessità potrebbe sorgere dal fatto che non abbiamo introdotto un calcolo. Abbiamo descritto tuttavia un algoritmo, che è basato solo sulla struttura sintattica, anche se infiorettato di terminologia semantica. Le risposte si evincono dalla struttura dell'albero (chiuso, non chiuso).

Piuttosto qualche ambiguità potrebbe sussistere in quanto le domande possibili sono diverse, ancorché collegate. Per il fatto che

Osservazione 6.1.3. Per ogni A,

$$A \text{ è una tautologia se e solo se } \neg A \text{ è insoddisfacibile,}$$

ci si può porre come problema semantico sia il problema della verità logica sia il problema dell'insoddisfacibilità. Un calcolo si può pensare sia come calcolo per stabilire la verità logica sia come un calcolo per stabilire l'insoddisfacibilità. Scegliamo il metodo degli alberi di refutazione per il problema dell'insoddisfacibilità, e come risposta preferenziale affermativa la chiusura dell'albero (un esito in generale più rapido e che non richiede ulteriori elaborazioni); abbiamo allora

Teorema 6.1.4 (Correttezza)**.** *Se l'albero di refutazione con radice A si chiude, allora A è insoddisfacibile.*

Dimostrazione. [2] Procediamo per contrapposizione dimostrando che se esiste un'interpretazione i che soddisfa A, allora a ogni stadio di sviluppo dell'albero esiste almeno un ramo tale che i soddisfa tutti le proposizioni del ramo. Allora l'albero non è mai chiuso, perché se un ramo è chiuso non tutte le sue proposizioni possono essere vere in una stessa interpretazione.

Allo stadio n, consideriamo un ramo σ le cui proposizioni siano tutte soddisfatte da i, e una proposizione B su di esso, quindi vera in i, e non ancora considerata (se non ce ne sono, il lavoro su quel ramo è terminato senza che esso sia chiuso, e tale rimane alla fine, e l'albero finale non è chiuso). Se B è una congiunzione, al ramo sono aggiunti due nodi che sono anch'essi etichettati con proposizioni vere in i, e il ramo prolungato soddisfa, allo stadio successivo, la proprietà richiesta. Se B è una disgiunzione $B_1 \vee B_2$, o il ramo[3] $\sigma^\frown B_1$ o il ramo $\sigma^\frown B_2$ soddisfano la proprietà richiesta, a seconda che B_1 o B_2 siano vere in i. Lo stesso vale per gli altri casi (esercizio). □

Viceversa

Teorema 6.1.5 (Completezza). *Se A è insoddisfacibile, l'albero di refutazione con radice p si chiude.*

Dimostrazione. Dimostriamo che

Lemma 6.1.6. *Se l'albero non si chiude, allora per ogni ramo non chiuso e terminato esiste un'interpretazione i che soddisfa tutti le proposizioni del ramo, inclusa la radice.*

Dimostrazione del Lemma. Sia σ un ramo non chiuso dell'albero terminato. Si definisca un'interpretazione i ponendo $i(p) = 1$ per ogni proposizione atomica p che occorre come nodo nel ramo σ, e $i(p) = 0$ per ogni proposizione atomica tale che $\neg p$ occorre come nodo nel ramo σ. Si dimostra ora che ogni proposizione di σ è vera in i. Supponiamo questo verificato per tutte le proposizioni sul ramo che hanno un'altezza minore di un numero fissato n, e facciamo vedere che lo stesso vale per quelle di altezza n. Se B è una congiunzione $B_1 \wedge B_2$, quando è stata presa in considerazione B si sono aggiunti come nodi del ramo sia B_1 che B_2, che sono quindi in σ e hanno altezza minore di n e quindi si suppongono vere in i; dunque anche B è vera in i. Se B è una disgiunzione $B_1 \vee B_2$, quando è stata presa in considerazione B si sono aggiunti a tutti i rami passanti per B, incluso (quello che sarebbe diventato) σ, o B_1 o B_2; quindi una delle due è su σ, e vera in i, quindi anche B è vera. Gli altri casi si trattano nello stesso modo. □

[2] Per questo e per il successivo teorema diamo dimostrazioni complete, impostate per induzione.

[3] $\sigma^\frown B_1$ è il ramo prolungato con B_1; la notazione è quella della concatenazione di liste.

$\square$

Se in un ramo terminato non chiuso manca una lettera che occorre nella radice, nel definire l'interpretazione si può dare ad essa il valore che si vuole; ciò significa che al ramo è associata più di una interpretazione.

L'esito complessivo dei teoremi di correttezza e completezza è che il metodo degli alberi prende in esame *tutte* le possibili strade per provare a definire interpretazioni, e se ce ne sono le fornisce tutte, e se non ce ne sono lo rivela.

La dimostrazione delle proprietà di correttezza e completezza non prende in considerazione l'ordine in cui si sviluppa l'albero. Il procedimento degli alberi di refutazione si può rendere deterministico fissando un ordine progressivo per le proposizioni introdotte e quelle da prendere in considerazione ma proprio il fatto che la dimostrazione è indipendente dall'ordine permette di vedere che la risposta dell'albero e le sue proprietà non dipendono dall'ordine eventualmente fissato; lavorare su una proposizione prima che su di un'altra può modificare l'albero ma non la risposta finale; ogni mossa dipende solo dalla proposizione in considerazione e non dalle altre presenti in altri nodi.

Si può sfruttare questa circostanza (oltre che come si è fatto nella dimostrazione della terminazione) per formulare utili regole euristiche, come quella di prendere in esame prima le proposizioni che si limitano ad allungare i rami e non introducono diramazioni.

Riassumendo

Corollario 6.1.7. *Per ogni A,*

A è soddisfacibile se e solo se l'albero di refutazione con radice A non si chiude.

mentre, nello spirito del controesempio,

Corollario 6.1.8. *Per ogni A,*

A è una tautologia se e solo se l'albero di refutazione con radice $\neg A$ si chiude.

Per la nozione di conseguenza logica, serve infine il

Corollario 6.1.9. *Per ogni A e B,*

$$\models A \to B$$

se e solo se l'albero di refutazione con radice $\neg(A \to B)$, o con radice $A \wedge \neg B$, si chiude.

Si noti che è indifferente avere nella radice $\neg(A \to B)$ oppure l'equivalente $A \wedge \neg B$ perché in entrambi i casi l'applicazione delle regole per la negazione di un condizionale o per la congiunzione portano ad aggiungere alla radice

$$\begin{array}{c} \downarrow \\ A \\ \downarrow \\ \neg B \end{array}$$

dopo di che si continua lavorando solo su A e su $\neg B$ e loro sottoproposizioni. Si può addirittura partire con

$$\begin{array}{c} A \\ \downarrow \\ \neg B \end{array}$$

se interessa la domanda $A \models B$.

Tuttavia, salvo questa eccezione iniziale, il metodo non prevede, e così le prove di esame, che nel corso della costruzione dell'albero una proposizione venga rimpiazzata da una logicamente equivalente; nei nodi dell'albero devono comparire solo sottoproposizioni o negazioni di sottoproposizioni della radice.

6.1.2 Forme normali

Gli alberi di refutazione permettono di ottenere altre informazioni sulle proposizioni a cui si applicano. Se A è una proposizione soddisfacibile, e quindi l'albero di refutazione con radice A non si chiude, una forma normale disgiuntiva di A si può ottenere nel seguente modo: per ogni ramo terminato e non chiuso, si faccia la congiunzione di tutti i letterali che sono nodi del ramo, quindi si faccia la disgiunzione di queste congiunzioni. Le proprietà dimostrate della correttezza e della completezza garantiscono che questa disgiunzione è proprio equivalente a A (esercizio).

Esempio

$$
\begin{array}{ccc}
\multicolumn{3}{c}{\neg(p \vee \neg q) \vee q \vee \neg(p \to q)} \\
\swarrow & \downarrow & \searrow \\
\neg(p \vee \neg q) & q & \neg(p \to q) \\
\downarrow & & \downarrow \\
\neg p & & p \\
\downarrow & & \downarrow \\
\neg\neg q & & \neg q \\
\downarrow & & \\
q\ . & &
\end{array}
$$

L'albero non è chiuso e la forma normale disgiuntiva della radice è $(\neg p \wedge q) \vee q \vee (p \wedge \neg q)$; i tre modelli dati dai tre rami non chiusi sono

$$
\begin{array}{l}
i_1(p) = 0, \quad i_1(q) = 1, \\
i_2(q) = 1, \\
i_3(p) = 1, \quad i_3(q) = 0
\end{array}
$$

dove il secondo sta per due interpretazioni, di cui una però coincide con la prima; rami diversi non danno interpretazioni diverse. La proposizione non è una tautologia in quanto manca l'interpretazione $i(p) = i(q) = 0$ tra i suoi modelli.

Se l'albero per A si chiude, si sa che A è una contraddizione e una forma normale disgiuntiva si scrive direttamente.

Dall'albero di A non si legge invece direttamente la forma normale congiuntiva di A; per ottenere questa, una via indiretta ma comunque veloce è la seguente: si mette nella radice $\neg A$, si sviluppa l'albero per $\neg A$ e si trova una forma normale disgiuntiva di $\neg A$. Quindi si nega questa premettendo una negazione, e si applicano le leggi di De Morgan.

Poiché l'albero terminato e non chiuso permette di leggere i modelli della radice, per verificare che A è una tautologia si può anche sviluppare l'albero con radice A, e controllare che ci siano alla fine 2^n interpretazioni associate ai rami non chiusi, se A ha n lettere. Ma se la domanda è se A sia una tautologia, è più conveniente impostare l'albero con $\neg A$, perché se la risposta è positiva essa arriva dalla chiusura dell'albero, in generale più in fretta dello sviluppo integrale dell'albero con radice A e del conteggio dei modelli.

Esercizio 6.1.10. Verificare con gli alberi di refutazione le leggi logiche della sezione 2.2.1.

Esercizio 6.1.11. Verificare con gli alberi di refutazione se le seguenti proposizioni sono tautologie, e se no indicare i controesempi:

$$
\begin{aligned}
&(p \lor q) \land (r \to \neg p) \to (r \to q)\\
&((p \to \neg p) \land (q \to p)) \to \neg q\\
&(p \land q) \lor (\neg p \land q) \lor q\\
&(p \land q) \lor (\neg p \land q) \lor \neg q.
\end{aligned}
$$

Esercizio 6.1.12. Verificare con gli alberi di refutazione se le seguenti proposizioni sono insoddisfacibili:

$$
\begin{aligned}
&((p \lor q) \land (\neg p \lor q) \land \neg q) \to q\\
&(p \to \neg q) \land (\neg p \lor q)\\
&(p \to \neg q) \land \neg p \land q\\
&(p \to \neg q) \land p \land q\\
&(p \lor \neg q \lor r) \land \neg r \land (\neg p \lor q) \land \neg p.
\end{aligned}
$$

Esercizio 6.1.13. Trovare con gli alberi di refutazione la forma normale disgiuntiva e i modelli delle seguenti proposizioni:

$$
\begin{aligned}
&p \land q \to (p \to q)\\
&p \land q \to (p \to q \land r)\\
&(p \to (q \lor (p \land r))) \land (\neg p \land (q \to p)).
\end{aligned}
$$

Con gli alberi di refutazione trovare la forma normale congiuntiva delle seguenti proposizioni:

$$
\begin{aligned}
&p \land q \to (p \to q \land r)\\
&(p \lor q \to r) \land \neg p \to (p \lor r)\\
&(p \to (q \lor (p \land r))) \land (\neg p \land (q \to p)).
\end{aligned}
$$

Capitolo 7

Linguaggi predicativi

7.1 Sintassi

Finora abbiamo considerato formule contenenti variabili solo nell'algebra degli insiemi, e si trattava di formule prive di quantificatori[1], usate per definire i valori di operazioni su insiemi.

La logica proposizionale ha permesso di introdurre come in laboratorio, su un *toy model*, alcuni concetti, come analisi sintattica, interpretazione, modello, conseguenza logica, validità, e alcuni sistemi per svolgere deduzioni e rispondere alle domande logiche.

Ora rivediamo gli stessi concetti in riferimento ai linguaggi predicativi, che sono quelli che permettono di formalizzare tutto il discorso matematico, e invero ogni discorso.

7.1.1 Alfabeto

L'alfabeto è stato già descritto praticamente nel capitolo 1. Useremo le notazioni lì indicate. Oltre ai connettivi e alle parentesi, si devono avere nell'alfabeto le variabili, con i due quantificatori, e simboli di predicato, di funzione e dicostante.

L'unica precisazione da aggiungere, rispetto a quanto era prima lasciato implicito, è che ogni simbolo di predicato ha associato un numero intero $n \geq 1$ che indica il numero di argomenti, o di posti (e se $n > 1$ il simbolo è detto anche simbolo di relazione) e analogamente ogni simbolo di funzione ha associato un numero $n \geq 1$ che indica il numero di argomenti[2]. La proprietà

[1] Salvo nel caso della definizione dell'unione e dell'intersezione generalizzate, in 4.1, dove intervengono i quantificatori, ad esempio $\bigcup_{i=1}^{n} A_i = \{x \in U \mid \exists i(1 \leq i \leq n \land x \in A_i\})$.

[2] Nella scrittura non è necessario che n sia una parte del simbolo, come potrebbe essere P^n; l'importante è che lo si dica o sia chiaro dal contesto.

di essere un simbolo dell'alfabeto, e di quale categoria, deve essere decidibile, e nel caso di simboli di predicato o di funzione il numero di argomenti deve essere effettivamente dato.

Le variabili[3] sono disponibili in quantità illimitata, anche se ogni volta se ne utilizzeranno solo un numero finito[4].

La dotazione di simboli di predicato, di funzione e di costante, che costituiscono la parte extralogica dell'alfabeto di un linguaggio, differisce da linguaggio a linguaggio, e possono anche mancare, anche se almeno un simbolo di predicato deve sempre essere presente.

Con "linguaggio" s'intende talvolta il complesso di alfabeto, espressioni, regole sintattiche e nozioni semantiche, altre volte semplicemente l'insieme delle formule (o delle proposizioni nel caso del linguaggio proposizionale).

7.1.2 Termini e formule

La struttura di base di un'affermazione atomica, come abbiamo visto, è l'attribuzione di un predicato a uno o più termini. Se $t_1, \ldots, t_n$ sono termini, non necessariamente distinti[5], si scriverà

$$P(t_1, \ldots, t_n)$$

a indicare che il predicato P sussiste per gli individui denotati dagli n termini (più precisamente la proprietà P se il numero di argomenti di P è $n = 1$, o la relazione P se $n > 1$).

I termini sono le costanti, le variabili e, se f è un simbolo di funzione a n posti, e $t_1, \ldots, t_n$ sono n termini, non necessariamente distinti, $f(t_1, \ldots, t_n)$ è un termine.

La scrittura $f(t_1, \ldots, t_n)$ è quella usuale per indicare il valore di f in corrispondenza agli argomenti $t_1, \ldots, t_n$. Ma le virgole non appartengono al linguaggio predicativo; le parole sono liste di simboli dell'alfabeto e quella che costituisce in effetti il termine in questione è semplicemente la concatenazione $ft_1 \ldots t_n$[6].

---

[3]Talvolta sono chiamate variabili individuali, per distinguerle da variabili per predicati o altri enti più astratti che si trovano in altri tipi di linguaggi, che non presentiamo. L'aggettivo "individuale" si riferisce al fatto che denotano *elementi* del dominio di discorso.

[4]Per alcune considerazioni, serve che l'insieme delle variabili sia ordinato, $x_1, x_2, \ldots$, e la funzione che a ogni n associa l'n-esima variabile sia una funzione calcolabile.

[5]Si potrebbe anche dire soltanto: "data una n-upla di termini", dove è implicito che le componenti non sono necessariamente distinte.

[6]$ft_1 \ldots t_n$ non è in generale una $(n+1)$-upla, ma una lista la cui lunghezza è $1+\sum_1^n l(t_i)$, dove $l(t_i)$ è la lunghezza della lista t_i: $ft_1 \ldots t_n$ è in effetti $f^\frown t_1^\frown \ldots ^\frown t_n$.

Si può dimostrare che non esiste ambiguità nel ricostruire come è composto il termine a partire da quali argomenti e quale simbolo funzionale, anche senza parentesi (e tanto meno virgole).

Questo fatto mostra i vantaggi della notazione prefissa rispetto a quella infissa nella manipolazione meccanica.

Lo stesso vale per la scrittura $P(t_1, \ldots, t_n)$, che useremo ma che corrisponde alla parola $Pt_1 \ldots t_n$.

Quando si trattano argomenti matematici, si usano le convenzioni a cui si è abituati, di scrivere i simboli delle operazioni usuali (che sono funzioni) in mezzo ai termini, laddove la notazione funzionale preferisce mettere il simbolo di funzione davanti; la stessa notazione infissa si adotta per le relazioni $=$, $<$ e $\leq$.

I termini *chiusi* sono i termini che non contengono variabili. I termini sono sempre infiniti, ma anchei termini chiusi lo sono se oltre a una costante c'è un simbolo funzionale.

Esempio 7.1.1. Supponiamo di avere una costante 0 e un simbolo funzionale a un argomento, indicato con s. I termini possono essere enumerati in una successione, ad esempio

$$0, s0, x, ss0, sx, y, sss0, ssx, sy, z, \ldots$$

Il criterio che guida l'enumerazione è quello, dopo il primo passo iniziale consistente nel porre 0 e $s0$, di introdurre una nuova variabile, aggiungere un s ai termini precedenti, e ricominciare con una nuova variabile.

I termini chiusi sono invece

$$0, s0, ss0, sss0, \ldots$$

Le versioni formali delle frasi saranno chiamate formule, in analogia alle formule matematiche.

Le formule sono definite nel seguente modo:

- Se P è un predicato a n posti e $t_1, \ldots, t_n$ termini, $(P(t_1, \ldots, t_n))$ è una formula.

- Se A è una formula, anche $(\neg A)$ lo è.

- Se A e B sono formule e $\bullet$ un connettivo binario, anche $(A \bullet B)$ è una formula.

- Se A è una formula, e x una variabile, anche $(\forall x A)$ e $(\exists x A)$ sono formule.

Le parentesi sono necessarie per riconoscere la struttura formale, il segno logico principale e le sottoformule. Le parentesi si riducono con le stesse convenzioni viste per le proposizioni, dove ora i quantificatori sono al primo posto nell'ordine di priorità, insieme alla negazione (se adiacenti, si procede alla reintroduzione dall'interno, o da destra verso sinistra, come già per la negazione nel linguaggio proposizionale).

Esempio 7.1.2. $\exists x P(x) \land \exists y \neg Q(y) \to \neg \exists x \exists y R(x,y)$ rimettendo le parentesi, escluse quelle intorno alle formule atomiche e la coppia esterna, diventa

$$\begin{gathered}
\exists x P(x) \land \exists y \neg Q(y) \to \neg \exists x (\exists y R(x,y)) \\
\exists x P(x) \land \exists y \neg Q(y) \to \neg (\exists x (\exists y R(x,y))) \\
\exists x P(x) \land \exists y \neg Q(y) \to (\neg (\exists x (\exists y R(x,y)))) \\
\exists x P(x) \land \exists y (\neg Q(y)) \to (\neg (\exists x (\exists y R(x,y)))) \\
\exists x P(x) \land (\exists y (\neg Q(y))) \to (\neg (\exists x (\exists y R(x,y)))) \\
(\exists x P(x)) \land (\exists y (\neg Q(y))) \to (\neg (\exists x (\exists y R(x,y)))) \\
((\exists x P(x)) \land (\exists y (\neg Q(y)))) \to (\neg (\exists x (\exists y R(x,y))))
\end{gathered}$$

che è una selva di parentesi ma da cui si vede la struttura, che è quella di un condizionale con l'antecedente che è una congiunzione e il conseguente che è una negazione; $\land$ e $\neg$ collegano tra loro formule quantificate (cioè che iniziano con un quantificatore).

Per le formule si possono costruire gli alberi sintattici individuando il segno logico principale, che ora può essere un connettivo oppure un quantificatore, per le formule del tipo $(\forall x A)$ e $(\exists x A)$.

Nei nodi dell'albero sintattico di una formula occorrono le sottoformule della formula stessa.

L'albero sintattico per $\exists x P(x) \land \exists y \neg Q(y) \to \neg \exists x \exists y R(x,y)$ è il seguente:

$$\begin{array}{ccc}
 & \exists x P(x) \land \exists y \neg Q(y) \to \neg \exists x \exists y R(x,y) & \\
 & \swarrow \quad \searrow & \\
\exists x P(x) \land \exists y \neg Q(y) & & \neg \exists x \exists y R(x,y) \\
\swarrow \quad \searrow & & \downarrow \\
\exists x P(x) \qquad \exists y \neg Q(y) & & \exists x \exists y R(x,y) \\
\downarrow \qquad\qquad \downarrow & & \downarrow \\
P(x) \qquad\quad \neg Q(y) & & \exists y R(x,y) \\
\qquad\qquad \downarrow & & \downarrow \\
\qquad\qquad Q(y) & & R(x,y).
\end{array}$$

Esempio 7.1.3. Le identità booleane sono esempi di formule atomiche.

Esempio 7.1.4. Tipiche formule atomiche di argomenti aritmetici e algebrici sono $x - 1 = 0$, $x \cdot y + x = x \cdot (y + 1)$, $x + (y + z) = (x + y + z)$, $x < x + 1$, e in generale $t_1 = t_2$ e $t_1 \leq t_2$ o $t_1 < t_2$ dove t_1 e t_2 sono termini.

Esempio 7.1.5. Tutti gli esempi del paragrafo 2.5.2 sono esempi di formule predicative.

7.1.3 Variabili libere e vincolate

La clausola 4 della definizione del linguaggio predicativo è la clausola nuova, rispetto al linguaggio proposizionale.

Per semplificare la discussione seguente, scriviamo Qx per indicare indifferentemente $\forall x$ o $\exists x$, nelle considerazioni che valgono per entrambi.

In una formula del tipo (QxA), A si chiama *raggio d'azione* del quantificatore Qx iniziale. Tutte le occorrenze di x all'interno del raggio d'azione di (Qx) (con una riserva che vedremo) vanno intese in senso o universale o rispettivamente esistenziale, a seconda di cosa è Q.

Naturalmente in (QxA) fuori dal raggio d'azione del primo quantificatore non c'è nulla, ma si deve pensare a quando (QxA) è parte di un'altra formula, ad esempio $(QxA) \wedge B$.

Per le occorrenze di x al di fuori del raggio d'azione del quantificatore, se non cadono dentro al raggio d'azione di un altro quantificatore, il senso in cui vanno interpretate non è determinato. L'interpretazione di tutta la formula allora è ambigua, o necessita di ulteriori precisazioni per essere compresa.

Ad esempio $\forall x(x^2 + 1 > 0)$ ha un senso compiuto, è l'affermazione di un fatto, e in particolare è vera in tutti i domini numerici usuali ordinati; nella formula

$$(\forall x(x^2 + 1 > 0)) \wedge x < 0$$

invece l'ultima x è indeterminata, non cadendo nel raggio d'azione di nessun quantificatore[7]. Si può studiare l'insieme di verità associato; nell'universo dei reali, per esempio, tale insieme è l'insieme dei numeri negativi; nell'universo dei naturali, è vuoto. La prima parte $\forall x(x^2 + 1 > 0)$ della congiunzione è vera in entrambi i casi e non contribuisce nulla alla delimitazione dell'insieme di verità.

In $\forall x(x^2 + 1 > 0) \wedge \exists x(x < 0)$ l'ultima occorrenza di x cade nel raggio d'azione di $\exists x$, e si ha la congiunzione di due formule che corrispondono entrambe ad affermazioni di senso compiuto, entrambe vere negli interi, razionali o reali, mentre la seconda è falsa nei naturali.

[7] Abbiamo messo ancora qui le parentesi esterne in $(\forall x(x^2 + 1 > 0))$ perché fosse chiaro dove finisce il raggio d'azione del quantificatore universale.

Le occorrenze di una variabile entro il raggio d'azione di un quantificatore relativo a quella variabile si dicono *vincolate* dal quantificatore nella formula (e così pure la x adiacente a Q in Qx, che spesso non viene neanche menzionata); quelle che non occorrono entro il raggio d'azione di un quantificatore relativo a quella variabile si dicono *libere*.

Se si vuole mettere in evidenza che la formula A contiene occorrenze libere di x si scrive $A[x]$, senza escludere che ne contenga altre; se si vuole mettere in evidenza che contiene occorrenze libere di x e di y scriviamo $A[x, y]$.

Qualche volta si dice brevemente che x è libera in A per dire che in A vi sono occorrenze libere di x, o che x è vincolata per dire che vi sono occorrenze vincolate, ma bisogna fare attenzione che allora come abbiamo visto una variabile può essere sia libera sia vincolata in una formula.

Le formule in cui non ci sono occorrenze libere di variabili si dicono *enunciati*. Sono le formule per cui ha senso chiedere se sono vere o false (una volta fissata l'interpretazione con il dominio di discorso).

Le formule che non sono enunciati saranno anche chiamate, per sottolineare la differenza, *formule aperte*; esse non esprimono frasi, piuttosto definiscono insiemi, o relazioni, a seconda di quante variabili libere hanno.

Esempio 7.1.6. In $\forall x(x^2 + 1 > 0) \land x < 0$ le prime due occorrenze di x sono vincolate; la terza è libera.

In $\forall x(x^2 + 1 > 0) \land \exists x(x < 0)$ tutte le occorrenze della x sono vincolate, ma le prime due dal quantificatore universale, le altre dal quantificatore esistenziale.

Un quantificatore può cadere entro il raggio d'azione di un altro quantificatore, come si è visto in diversi esempi.

Ma un quantificatore relativo a una variabile x può anche cadere entro il raggio d'azione di un altro quantificatore relativo alla stessa x. Ad esempio, dopo aver considerato la formula del tipo $A[x]$, con una variabile x che occorre libera:

$$\forall x(x^2 + 1 > 0) \land x < 0$$

nel dominio degli interi, e aver verificato che il suo insieme di verità non è vuoto, si ottiene un enunciato vero premettendo $\exists x$. Ma allora si ottiene l'enunciato

$$\exists x(\forall x(x^2 + 1 > 0) \land x < 0),$$

che richiede di essere letto con attenzione. Quando nella costruzione di una formula si premette ad A un quantificatore con la variabile x, questo quantificatore vincola tutte le occorrenze di x che sono libere in A, e solo quelle. Proprio per come è stato ottenuto, è chiaro che il quantificatore esistenziale nell'esempio vincola l'occorrenza di x che prima era libera, cioè l'ultima, e solo quella. L'azione del quantificatore esistenziale premesso $\exists x$ scavalca la parte $\lceil \ldots \rceil$ in

$$\lceil \forall x(x^2 + 1 > 0) \land \rceil x < 0,$$

dove non ci sono occorrenze libere di x, per agire su $x < 0$ dove x occorre libera. Le occorrenze vincolate di x in $\lceil \ldots \rceil$, essendo già vincolate, sono insensibili all'azione di un altro quantificatore. In effetti, è come se fosse scritto ad esempio

$$\forall z(z^2 + 1 > 0) \land x < 0,$$

e si ottenesse perciò

$$\exists x(\forall z(z^2 + 1 > 0) \land x < 0),$$

che è un modo di scrivere l'enunciato più chiaro, ed equivalente. Se si legge la frase in italiano si vede bene che non c'è interferenza tra le occorrenze libere e vincolate di x, perché si possono usare locuzioni diverse; "esiste un numero tale che, mentre ogni numero elevato al quadrato e aumentato di 1 è maggiore di 0, lui è negativo"[8]. Ancor meglio, conviene leggere: "mentre ogni numero elevato al quadrato e aumentato di 1 è maggiore di 0, esiste un numero che è negativo". Infatti un altro modo di evitare difficoltà interpretative è quello di andare a piazzare il nuovo $\exists x$ dove è richiesto, cioè scrivendo

$$\forall x(x^2 + 1 > 0) \land \exists x(x < 0).$$

Vedremo in seguito che tali trasformazioni equivalenti sono legittime.

Infine un quantificatore relativo ad una variabile x si può premettere anche a una formula che non contenga alcuna occorrenza di x libera, ad esempio $\exists x \forall y(y^2 + 1 > 0)$ o $\forall x(y < 0)$. La definizione di "formula" non lo esclude[9]. In questi casi l'effetto del primo quantificatore è nullo, la sua presenza superflua, e la formula ottenuta equivalente a quella originaria.

7.2 Semantica

Le formule matematiche presentate negli esempi visti finora possono essere interpretate in diversi domini numerici; alcune sono vere negli uni e false negli altri. La possibilità di diverse interpretazioni è ancora più evidente in formule del tipo $\forall x(\exists y A(x, y) \to F(x))$, dove ci sono simboli predicativi A

[8] Si è usato qui "mentre" come congiunzione, per sottolineare la non connessione tra le due parti della frase.

[9] Non lo esclude perché sarebbe stato complicato inserire la condizione sulle occorrenze libere nella definizione stessa iniziale.

e F che non hanno un'interpretazione nemmeno nell'uso comune (come è il caso dei simboli matematici), e questa è la caratteristica della logica *formale*.

$\forall x(\exists y A(x, y) \to F(x))$ ad esempio può essere interpretato:

- nell'universo delle persone, e significare che chi ha un amico è felice — se A e F sono interpretati in questo modo,
- o significare che uno è felice se ha un lavoro, se A è interpretato in quest'altro modo,
- o significare che il dominio della relazione R è contenuto in X, se la relazione R e l'insieme X sono definiti rispettivamente da $A(x, y)$ e $F(x)$,
- oppure essere interpretato nei numeri naturali, usando ad esempio $A(x, y)$ per "x è divisibile per y con quoziente maggiore di 1 e minore di x" e $F(x)$ per "x è un numero composto", e l'enunciato è vero in questa interpretazione.

Prima di chiedersi se un enunciato è vero o no occorre precisare quale interpretazione si ha in mente. Tuttavia per lo studio generale delle interpretazioni, le singole interpretazioni sono troppo varie. Come nel caso proposizionale, si cerca di formulare una definizione matematica astratta del concetto di interpretazione.

7.2.1 Interpretazioni

Ogni interpretazione effettiva in un dominio di conoscenze comporta che si individui l'universo del discorso, che deve essere un insieme non vuoto; quindi si devono stabilire le relazioni e funzioni su questo insieme che corrispondono ai simboli predicativi e funzionali che occorrono nell'enunciato. Se ci sono costanti, bisogna fissare gli elementi di U di cui le costanti sono nomi.

Una interpretazione per un linguaggio predicativo è allora una *struttura*

$$\mathcal{M} = \langle M, R^{\mathcal{M}}, \dots, f^{\mathcal{M}}, \dots, c^{\mathcal{M}}, \dots \rangle$$

dove M è un insieme non vuoto, $R^{\mathcal{M}}, \dots$ sono relazioni n-arie su M, per ogni simbolo relazionale R a n argomenti dell'alfabeto, $f^{\mathcal{M}}, \dots$ sono funzioni da M^n in M, per ogni simbolo funzionale f dell'alfabeto, e $c^{\mathcal{M}}, \dots$ sono elementi di M, uno per ogni simbolo di costante.

Data un'interpretazione, solo alle variabili, di tutti i simboli dell'alfabeto, manca un riferimento nellla struttura. Per dare una denotazione a tutti i termini e un senso alle formule aperte occorre fissare il riferimento delle variabili.

Si dice *assegnazione* in $\mathcal{M}$ una funzione α dall'insieme delle variabili in M. Fissata un'assegnazione, a ogni termine t viene a corrispondere un elemento t^α di M secondo la seguente definizione induttiva:

- Se t è una variabile x, t^α è $\alpha(x)$
- Se t è una costante c, t^α è $c^\mathcal{M}$
- Se t è $f(t_1, \ldots, t_n)$, t^α è $f^\mathcal{M}(t_1^\alpha, \ldots, t_n^\alpha)$.

Se t è un termine chiuso, t^α è uguale per ogni α, e sarà indicato con $t^\mathcal{M}$.

Fissata un'assegnazione α in $\mathcal{M}$, a ogni formula corrisponde un valore di verità nell'interpretazione $\mathcal{M}$, sotto l'assegnazione. Si dirà che α *soddisfa* A *in* $\mathcal{M}$, e si scriverà

$$\mathcal{M}, \alpha \models A.$$

Per determinare tale valore, si parte dalle sottoformule atomiche, per cui

$$\mathcal{M}, \alpha \models P(t_1, \ldots, t_n) \text{ se e solo se } \langle t_1^\alpha, \ldots, t_n^\alpha \rangle \in P^\mathcal{M}$$

e si risale usando le tavole di verità per i connettivi, e la naturale interpretazione per i quantificatori.

Si vede facilmente che se le assegnazioni α e β coincidono sull'insieme delle variabili occorrenti in t, allora $t^\alpha = t^\beta$; e che se α e β coincidono sull'insieme delle variabili libere di A, allora $\mathcal{M}, \alpha \models A$ se e solo se $\mathcal{M}, \beta \models A$.

Se A è un enunciato quindi, due qualunque assegnazioni coincidono sull'insieme (vuoto) delle variabili libere di A, e quindi o tutte le assegnazioni lo soddisfano, e in tal caso si dice che A è *vero* in $\mathcal{M}$, o che $\mathcal{M}$ è *modello* di A, e si scrive

$$\mathcal{M} \models A$$

oppure nessuna lo soddisfa, e in tal caso A si scrive $\mathcal{M} \not\models A$ e si dice che A è *falso* in $\mathcal{M}$[10].

Validità e conseguenza

Una formula A si dice *soddisfacibile in* $\mathcal{M}$ se esiste qualche assegnazione α tale che $\mathcal{M}, \alpha \models A$.

Una formula A si dice *valida* in $\mathcal{M}$, o che $\mathcal{M}$ è *modello* di A, e si scriverà $\mathcal{M} \models A$, se $\mathcal{M}, \alpha \models A$ per ogni α in $\mathcal{M}$.

---

[10] Perché la dicotomia sussista è necessario che esista almeno una assegnazione, quindi che M non sia vuoto.

Una formula A si dice *soddisfacibile* se esiste qualche $\mathcal{M}$ in cui è valida. Altrimenti si dice insoddisfacibile.

Se A è un enunciato, e come formula è valida in $\mathcal{M}$, si dice che A è vero in $\mathcal{M}$.

La *chiusura universale* di A è l'enunciato $\forall A$ che si ottiene premettendo ad A un quantificatore universale per ogni variabile libera in A. Una formula A è valida in $\mathcal{M}$ se e solo se $\forall A$ è vero in $\mathcal{M}$.

Si dirà che A è *logicamente valida*, e si scriverà $\models A$, se

$$\mathcal{M} \models A, \text{ per ogni } \mathcal{M}.$$

Se A è una formula chiusa logicamente valida, A si dirà anche un enunciato *logicamente vero*.

A si dice *conseguenza logica* di $A_1, \ldots, A_n$, e si scriverà

$$A_1, \ldots, A_n \models A$$

se per ogni $\mathcal{M}$ e ogni assegnazione α in $\mathcal{M}$,

$$\text{se } \mathcal{M}, \alpha \models A_1 \wedge \ldots \wedge A_n \text{ allora } \mathcal{M}, \alpha \models A,$$

ovvero se $\models A_1 \wedge \ldots \wedge A_n \to A$.

Se diciamo che $\mathcal{M}, \alpha \models T$ se $\mathcal{M}, \alpha \models B$ per ogni $B \in T$, la definizione di conseguenza $T \models A$ si può generalizzare a insiemi qualunque T.

Se T è finito $T = \{A_1, \ldots, A_n\}$, per ogni A si ha che $A_1, \ldots, A_n \models A$ se e solo se $A_1 \wedge \ldots \wedge A_n \models A$.

Se A e B sono enunciati, $B \models A$ se e solo se ogni modello di B è anche modello di A.

A è *logicamente equivalente* a B, in simboli $A \equiv B$ se per ogni $\mathcal{M}$ e per ogni assegnazione α in $\mathcal{M}$,

$$\mathcal{M}, \alpha \models A \text{ se e solo se } \mathcal{M}, \alpha \models B,$$

che corrisponde a $\models A \leftrightarrow B$.

La sostituzione

Discutiamo a parte alcuni problemi che si pongono per la sostituzione di termini alle variabili libere di una formula.

Indichiamo con $A[x/t]$ il risultato della sostituzione del termine t a tutte le occorrenze libere di x in t, senza escludere che in A ci siano altre variabili libere[11].

[11]In precedenza avevamo indicato con $A[t]$ il risultato della sostituzione, ma conviene tenere traccia della variabile sostituita.

Indichiamo con $s[x/s_1]$ il risultato della sostituzione del termine s_1 a tutte le occorrenze della variabile x in s.

Innanzi tutto osserviamo come può essere definita l'operazione di sostituzione. Essa infatti non consiste nella semplice scrittura del termine nel posto occupato dalla variabile, a meno che il termine non sia un'altra variabile o una costante, cioè di lunghezza 1. In generale il risultato è una espressione di diversa lunghezza.

Si definisce prima la sostituzione $s[x/s_1]$:

- se s è x, $s[x/s_1]$ è s_1;

- se s è una variabile diversa da x o una costante, $s[x/s_1]$ è s;

- se s è $f(t_1,\ldots,t_n)$, $s[x/s_1]$ è $f(t_1[x/s_1],\ldots,t_n[x/s_1])$.

Quindi si definisce la sostituzione $A[x/t_1]$:

- se A è $P(s_1,\ldots,s_n)$, $A[x/t]$ è $P(s_1[x/t],\ldots,s_n[x/t])$;

- se A è $\neg B$, $A[x/t]$ è $\neg(B[x/t])$;

- se A è $B \bullet C$, $A[x/t]$ è $B[x/t] \bullet C[x/t]$;

- se A è QxB, $A[x/t]$ è A;

- se A è QyB, y diversa da x, $A[x/t]$ è $Qy(B[x/t])$.

Quando si esegue una sostituzione $A[x/t]$, si pensa che l'affermazione fatta da A su x venga trasportata a t. Ma occorre cautela, possono presentarsi incongruenze dovute alla rigidità della sintassi.

Se ad esempio A è la formula $\exists y(x \neq y)$, allora A afferma intuitivamente che esiste un elemento diverso da x. Quindi

$$\begin{gathered}\exists y(0 \neq y) \text{ afferma che esiste un elemento diverso da } 0,\\ \exists y(1+1 \neq y) \text{ afferma che esiste un elemento diverso da } 1+1,\\ \exists y(2 \cdot z \neq y) \text{ afferma che esiste un elemento diverso da } 2 \cdot z\end{gathered}$$

e così via.

Le frasi ottenute per sostituzione conservano lo stesso senso, si leggono nello stesso modo.

Se però si sostituisce a x la variabile y allora la formula che si ottiene non afferma per y la stessa cosa affermata per x, o per 0, o 1+1 o $2 \cdot z$: si ottiene $\exists y(y \neq y)$, che a parte che è falsa, non ha più lo stesso significato.

Come altro esempio, posto che $x|y \leftrightarrow \exists z(y = x \cdot z)$ significa che x divide y, allora:

- $1|y$ significa che 1 divide y,
- $0|y$ significa che 0 divide y,
- $(2 \cdot x)|y$ significa che $2 \cdot x$ divide y,
- $u|y$ significa che u divide y,

ma

- $\exists z(y = z \cdot z)$ significherebbe che y è un quadrato.

Un esempio tratto, forzosamente, dal linguaggio comune potrebbe essere il seguente, dove si suppone di usare "Tizio", "Caio" e "Sempronio" come variabili: invece di dire che ognuno ha un padre, si dica "ogni Tizio ha un Caio per padre"; particolarizzando, si può dedurre che Giovanni ha un Caio per padre, che Maria ha un Caio per padre, ma non si può dedurre "Caio ha Caio per padre".

Per isolare i casi di sostituzione illecita si introduce la seguente

Definizione 7.2.1. Un termine t si dice *sostituibile*[12] *a x in A* se non ci sono in A occorrenze di x che cadono dentro al raggio d'azione di quantificatori Qy, dove y sono variabili occorrenti in t.

Con la definizione si vuole intendere che se si sostituisce a x in A un termine sostituibile a x in A, la frase $A[x/t]$ ha lo stesso significato per t di A.

I rapporti tra $A[x]$ e $A[x/t]$ sono regolati dal seguente

Lemma 7.2.2. *Se t è sostituibile a x in A e se α e α' sono due assegnazioni in $\mathcal{M}$ tali che $\alpha'(x) = t^\alpha$, e $\alpha'(y) = \alpha(y)$ per ogni altra eventuale variabile libera in A diversa da x, allora*

$$\mathcal{M}, \alpha' \models A \text{ se e solo se } \mathcal{M}, \alpha \models A[x/t].$$

Come caso particolare:

Corollario 7.2.3. *Se y è una variabile che non occorre in A, né libera né vincolata, e se α e α' sono due assegnazioni in $\mathcal{M}$ tali che*

$$\alpha'(z) = \begin{cases} \alpha(y) & \text{se } z \text{ è } x, \\ \alpha(z) & \text{se } z \text{ è diversa da } x, \end{cases}$$

allora

$$\mathcal{M}, \alpha' \models A \text{ se e solo se } \mathcal{M}, \alpha \models A[x/y].$$

[12] Nella letteratura si trova più spesso la dizione "libero per x in A".

Il corollario permette di utilizzare la stessa soluzione che si usa di fatto nel parlato per evitare le incongruenze del tipo sopra segnalato dovute alla sostituzione di un termine non sostituibile[13].

Se, in riferimento al primo esempio di sopra, si vuole dire che esiste un elemento diverso da y, si scrive ad esempio $\exists u(y \neq u)$, prendendo come espressione di "esiste un elemento diverso da x" la formula $\exists u(x \neq u)$ invece di $\exists y(x \neq y)$.

Analogamente, in riferimento al secondo esempio, se si deve affermare che $z|y$ si assume che la definizione di "x divide y" sia data dalla formula $\exists v(y = x \cdot v)$ e non da $\exists z(y = x \cdot z)$

Le due formule sono logicamente equivalenti, grazie al

Lemma 7.2.4. *Se y è una variabile che non occorre in A, allora*

$$QxA \equiv QyA[x/y].$$

L'ipotesi è necessaria, come mostra il controesempio di $\exists x \exists y R(x, y)$ che non è equivalente a $\exists y \exists y R(y, y)$.

La trasformazione permessa dal lemma di QxA nella equivalente $QyA[x/y]$ si chiama *rinomina* delle variabili vincolate.

La definizione di sostituzione comporta che se x non occorre libera in A o non occorre in s, allora $A[x/t]$ o $s[x/s_1]$ sono rispettivamente uguali a A o a s.

Ne segue che se x non occorre libera in A allora $QxA \equiv A$.

Esercizi

Esercizio 7.2.5. In un linguaggio aritmetico con costanti per le prime dieci cifre da 0 a 9, e le operazioni di somma, prodotto e potenza, spiegare quali sono i termini che denotano i numeri nella notazione posizionale in base dieci; quali sono i termini per 512, per 23, per 101?

Esercizio 7.2.6. Scrivere in linguaggio predicativo tutte le definizioni relative alle relazioni d'ordine (massimo, minimo, maggiorante, minorante, estremo superiore, . . .).

Esercizio 7.2.7. Quali sono gli insiemi di verità (cfr. sezione 3.1.1) in $\mathbb{N}$ di

$$\begin{gathered}
\exists y(2y = x \wedge \exists z(2z = y)) \\
\exists y(2y = x \wedge \exists y(2y = x)) \\
\exists y(xy = 2 \wedge \exists z(yz = 2)) \\
\exists y(xy = 2 \wedge \exists y(xy = 2))?
\end{gathered}$$

[13]Basterebbe il lemma, ma in pratica si utilizza la possibilità più semplice data dal corollario, di prendere una variabile nuova.

7.2.2 Leggi logiche

Le formule logicamente valide continuano a chiamarsi anche leggi logiche.

Esempi di leggi logiche in linguaggi predicativi si ottengono facilmente partendo da tautologie proposizionali e rimpiazzando le lettere che vi compaiono con formule qualunque del linguaggio, la stessa formula a tutte le occorrenze della stessa lettera.

Ad esempio da $\models p \vee \neg p$ segue $\models \exists x P(x) \vee \neg \exists x P(x)$. Infatti, data una qualunque intepretazione, con un qualunque predicato per P, in essa $\exists x P(x)$ risulterà o vero o falso. Se risulta vero, è come se si assegnasse il valore 1 a p nella proposizione; se risulta falso, è come se si assegnasse 0 a p nella proposizione; ma questa è una tautologia, per cui risulta vera in entrambi i casi, e i calcoli che si fanno a partire dai valori di p per arrivare al valore dalla proposizione $p \vee \neg p$ sono gli stessi che si fanno a partire dal fatto che $\exists x P(x)$ è vero o no per arrivare a dire se $\exists x P(x) \vee \neg \exists x P(x)$ è vero o no. Non c'è bisogno di considerare alcuna interpretazione e vedere se in essa $\exists x P(x)$ è vero o falso, perché comunque essa sia, e comunque sia l'interpretazione di P, e quindi il valore di $\exists x P(x)$, in essa $\exists x P(x) \vee \neg \exists x P(x)$ risulterà vero.

Lo stesso succede con qualsiasi altra tautologia, e con la sostituzione di una qualunque formula.

Quindi tutte le leggi logiche proposizionali restano tali considerando ora le lettere A, B, ... che vi compaiono come formule di un qualunque linguaggio predicativo.

Ma esistono anche altre leggi logiche per formule con quantificatori che non si ottengono in questo modo e dipendono proprio dal significato dei quantificatori. Ad esempio

$$\forall x \neg A \leftrightarrow \neg \exists x A$$

è una di queste[14].

Per verificarlo si ragiona nel seguente modo:

Dimostrazione. In una qualunque interpretazione, se $\forall x \neg A$ è vero, l'insieme di verità di $\neg A$ è tutto l'universo, quindi l'insieme di verità di A è vuoto; allora $\exists x A$ è falso, e quindi $\neg \exists x A$ è vero. □

La legge si può considerare una generalizzazione di quelle di De Morgan, se si pensa che affermare $\forall x A$ sia come fare una grande congiunzione per tutte le $A[x]$, al variare di x nell'universo, e affermare $\exists x A$, cioè che $A[x]$ vale per almeno un x, sia come fare una grande disgiunzione.

[14]Nella verifica di questa e delle successive leggi logiche, come già nella precedente della forma $A \vee \neg A$, supporremo per semplicità che si tratti di enunciati, solo per mostrare in modo più facile l'idea soggiacente.

Si è visto già nella definizione di unione e intersezione generalizzate come i quantificatori esistenziale ed universale siano usati come generalizzazione della disgiunzione e della congiunzione.

Se si combina questa legge logica con quella della doppia negazione si ottengono altre versioni, come

$$\neg\forall x\neg A \leftrightarrow \exists x A$$

o

$$\forall x A \leftrightarrow \neg\exists x\neg A$$

che mostrano come i due quantificatori non siano indipendenti, ma l'uno definibile in termini dell'altro, e della negazione. Si chiamano anche leggi della interdefinibilità dei quantificatori.

La legge tipica del quantificatore universale è la legge di *particolarizzazione*

$$\forall x A \to A[x/t],$$

per ogni termine sostituibile a x che A ha per x.

Le applicazioni della legge sono frequenti; gli assiomi di una teoria sono in genere enunciati che iniziano con un quantificatore universale (oppure sono presentati come formule valide, supponendo tacitamente una possibilità di sostituzione di termini qualsiasi alle variabili che è di fatto un'applicazione della particolarizzazione).

Si trovano sia esempi di applicazioni in cui t è chiuso sia esempi in cui contiene variabili.

Esempio 7.2.8. La legge boleana dell'unicità dell'elemento neutro dell'addizione

$$\forall x(x + y = x) \to y = 0$$

si può dimostrare in questi due modi.

Applicando a $\forall x(x+y = x)$ la particolarizzazione con 0 si ottiene $0+y = 0$ da cui con trasformazioni algebriche, utilizzando $y + 0 = y$, si arriva a $y = 0$.

Applicando invece a $\forall x(x+y = x)$ la particolarizzazione con $-y$ si ottiene $-y + y = -y$, quindi $1 = -y$ e $y = 0$.

Nell'esempio la formula quantificata universalmente è del tipo $\forall x A[x, y]$, e $-y$ è ovviamente sostituibile a x perché $A[x, y]$ non contiene quantificatori.

La validità della legge di particolarizzazione è una facile conseguenza del lemma 7.2.2.

Se t non è sostituibile a x in A, il condizionale non è valido, come mostra il controesempio $\forall x\exists y R(x, y) \to \exists y R(y, y)$, che è falso nella struttura dei numeri naturali se R è interpretato sulla relazione $<$.

Una legge simmetrica rispetto a quella di particolarizzazione, che si chiama anche di *generalizzazione esistenziale*, o di indebolimento esistenziale, afferma che se t è sostituibile a x in A allora è logicamente valida

$$A[x/t] \to \exists x A.$$

Ad esempio, siccome $y+(-y)=0$ vale negli interi, si può dedurre $\exists x(y+x=0)$; qui bisogna pensare che $Ax/t]$ è $y+(-y)=0$, con t uguale a $-y$, e che è ottenuta da $y+x=0$ per sostituzione di $-y$ a x. Ma potrebbe essere stata ottenuta per sostituzione di $-y$ a z in $y+z=0$ e si può altrettanto correttamente dedurre $\exists z(y+z=0)$, pensando di applicare $A[z/t] \to \exists z A$.

Se si combinano in serie particolarizzazione e generalizzazione esistenziale si ottiene

$$\forall x A \to \exists x A,$$

che è valida in quanto si considerano solo sempre interpretazioni in cui l'universo non è vuoto[15].

Sono leggi logiche anche le leggi di rinomina delle variabili vincolate del lemma 7.2.4.

Le leggi di rinomina seguono anche dalla legge di particolarizzazione e dalle prossime leggi riguardanti condizionale e quantificatori (esercizio).

Altre leggi stabiliscono dei rapporti tra connettivi e quantificatori che permettono di trasformare le formule in altre equivalenti con un diverso segno logico principale:

$$\begin{aligned} \forall x(A \land B) &\equiv \forall x A \land \forall x B \qquad && \text{distributività di } \forall \text{ su } \land \\ \exists x(A \lor B) &\equiv \exists x A \lor \exists x B \qquad && \text{distributività di } \exists \text{ su } \lor \end{aligned}$$

sono immediate conseguenze del significato dei simboli logici.

Si possono derivare facilmente con la deduzione naturale (esercizio).

Mentre è pure ovvio che siano logicamente valide

$$\forall x A \lor \forall x B \to \forall x(A \lor B)$$
$$\exists x(A \land B) \to \exists x A \land \exists x B,$$

non valgono le implicazioni inverse.

Se ad esempio U è un insieme con due elementi $\{a,b\}$ e l'interpretazione di P è $\{a\}$ e l'interpretazione di Q è $\{b\}$, allora in questa interpretazione $\forall x(P(x) \lor Q(x))$ è vero, mentre sono falsi sia $\forall x P(x)$ (l'insieme di verità di $P(x)$ non è tutto U) sia $\forall x Q(x)$.

[15]Se si scegliesse come U l'insieme di tutte le creature della fantasia, non si potrebbe pretendere, come non si pretende, che ivi valgano tutte le leggi della logica. Lo stesso nell'universo di quei filosofi che dicono che tutto è apparenza.

Esercizio. Si trovi un'interpretazione in cui $\exists x A \wedge \exists x B$ è vero e $\exists x(A \wedge B)$ è falso.

Sono particolarmente importanti le leggi che regolano i rapporti tra quantificatori e condizionale. Mentre è facile convincersi che

$$\forall x(A \to B) \to (\forall x A \to \forall x B)$$

è logicamente valida, il condizionale inverso non lo è. Non si può quindi parlare di distributività di $\forall$ su $\to$.

Per trovare un controesempio, si deve pensare ad un'interpretazione in cui $\forall x P(x) \to \forall x Q(x)$ sia vero semplicemente perché $\forall x P(x)$ è falso, mentre non è vero $\forall x(P(x) \to Q(x))$. L'insieme $U = \{a, b, c\}$ con $\{a, b\}$ per l'insieme di verità di $P(x)$ e $\{b, c\}$ per l'insieme di verità di $Q(x)$ risponde allo scopo.

Se A non contiene x libera tuttavia, allora

$$\forall x(A \to B) \equiv A \to \forall x B$$

e

$$\exists x(A \to B) \equiv A \to \exists x B.$$

Verifichiamo la prima.

Dimostrazione. Dato un U qualsiasi e supposto per semplicità che A sia un enunciato, distinguiamo due casi. Se A è falso, $A \to \forall x B$ è vero, ma d'altra parte anche $\forall x(A \to B)$ è vero perché $A \to B$ è soddisfatta da ogni assegnazione.

Se A è vero, l'insieme di verità di $A \to B$ è uguale all'insieme d i verità di B: le assegnazioni che soddisfano $A \to B$ sono esattamente quelle che soddisfano B. Quindi se $\forall x(A \to B)$ è vero, ma anche $\forall x B$ lo è, e così pure $A \to \forall x B$. $\square$

Queste leggi esprimono un caso particolare della possibilità di mettere il quantificatore nella posizione in cui il suo raggio d'azione esplica la sua funzione effettiva, sulle occorrenze libere della variabile (spostarlo all'indentro). Ne deriva una migliore leggibilità delle formule e una loro maggiore aderenza alle versioni informali.

Altri casi analoghi, e non indipendenti dai precedenti, sono le leggi

$$\forall x(A \vee B) \equiv A \vee \forall x B$$

e

$$\exists x(A \wedge B) \equiv A \wedge \exists x B$$

se x non occorre libera in A.

Ovviamente, per la commutatività di $\wedge$ e $\vee$, la formula senza la variabile libera può essere sia la prima sia la seconda della congiunzione e rispettivamente disgiunzione. Invece il condizionale non è commutativo. Tuttavia esistono leggi riguardanti i quantificatori nell'antecedente, forse a prima vista sorprendenti:

$$\forall x(A \to B) \equiv \exists x A \to B$$

e

$$\exists x(A \to B) \equiv \forall x A \to B$$

se x non occorre libera in B.

La prima legge corrisponde al seguente uso linguistico: quando si dice che qualche cosa, espressa da B, dipende solo dal fatto che (l'insieme di verità di) A non sia vuoto, e non da un particolare elemento (come sarebbe se x occorresse in B e dovesse soddisfarla), allora si può enfatizzare che *qualunque* elemento va bene. Se si afferma "se uno ha un amico è felice" — $\exists x A(x, y) \to F(y)$ — si vuol dire che qualunque sia l'amico, anche un cane, porta felicità (a y).

Dimostriamo la prima (l'altra è analoga)

Dimostrazione. Si supponga che l'enunciato (per semplicità) $\exists x A(x) \to B$ sia vero, quindi B vero o $\exists x A$ falso. Se B è vero, anche allora $A \to B$ è soddisfatta da ogni assegnazione, e $\forall x A \to B$ è vero. Se $\exists x A$ è falso, ogni assegnazione soddisfa $A \to B$, e $\forall x A \to B$ è vero.

Se invece $\exists x A \to B$ è falso, allora B è falso e $\exists x A$ è vero, quindi almeno una assegnazione soddisfa A. Questa assegnazione rende falso il condizionale $A \to B$, per cui $\forall x(A(x) \to B)$ non è vero. □

Esempio 7.2.9. Deduzione naturale di $\exists x A \to B \vdash \forall x(A \to B)$.

Supponiamo che le due formule siano enunciati.

1	$\exists x A \to B$	assunzione
2	A	assunzione da scaricare a 5
3	$\exists x A$	(I$\exists$) da 2
4	B	(MP) da 1 e 3
5	$A \to B$	(I$\to$) da 2 e 4
6	$\forall x(A \to B)$	(I$\forall$) da 5

Le equivalenze che permettono di spostare all'interno i quantificatori permettono anche di spostarli all'esterno; si ottiene così che ogni formula è equivalente ad una formula in cui tutti i quantificatori sono all'inizio — e formano il cosiddetto *prefisso* — seguiti da una formula senza quantificatori — detta *matrice*; una formula scritta in questo modo di dice in *forma prenessa*.

Se il prefisso è tutto costituito da quantificatori universali la formula si dice universale, se da quantificatori esistenziali si dice esistenziale.

Ad esempio

$$\forall x P(x) \vee \forall x Q(x) \to \forall x(P(x) \vee Q(x))$$

è equivalente (esercizio) a

$$\exists x \exists y \forall z(P(x) \vee Q(y) \to P(z) \vee Q(z))$$

ma anche a

$$\forall z \exists x \exists y(P(x) \vee Q(y) \to P(z) \vee Q(z))$$

ed esistono altre forme prenesse della stessa formula, a seconda dell'ordine in cui si esportano i quantificatori.

Si noti lo scambio dei quantificatori universali ed esistenziali nei due prefissi; lo scambio di solito non è permesso, non è valido in generale, ma capita con prefissi diversi di due forme premesse equivalenti di una stessa formula.

Le regole viste, ed elencate nella tabella 7.1 sono sufficienti a trasformare ogni formula in forma prenessa. Alcune ulteriori leggi utili si derivano da esse con l'uso della rinomina delle variabili vincolate. Ad esempio dalla distributività di $\forall$ su $\wedge$ segue

$$\forall x A \wedge B \equiv \forall x(A \wedge B) \qquad \text{se } x \text{ non libera in } B,$$

e analogamente dalla distributività di $\exists$ su $\vee$ segue

$$\exists x A \vee B \equiv \exists x(A \vee B) \qquad \text{se } x \text{ non libera in } B.$$

Per trasformare in forma prenessa $\forall x A[x] \wedge \exists x B[x]$ si può allora ottenere prima $\exists x(\forall x A[x] \wedge B[x])$; quindi poiché x è libera in B si sostituisce $\forall x A[x] \wedge B[x]$ con l'equivalente $\forall z(A[x/z] \wedge B(x))$, dopo una rinomina; si ottiene infine $\forall x A[x] \wedge \exists x B[x] \equiv \exists x \forall z(A[z] \wedge B[x])$, oppure in altro modo (esercizio) $\forall x A[x] \wedge \exists x B[x] \equiv \forall x \exists z(A[x] \wedge B[z])$.

Una formula si dice in *forma normale di Skolem* se la formula è universale e la sua matrice è in forma normale congiuntiva (dove ora con "letterale" si intende una formula atomica o la negazione di una formula atomica).

Due formule A e B si dicono *equisoddisfacibili* se A è soddisfacibile se e solo se B è soddisfacibile (non necessariamente con lo stesso modello).

Lemma 7.2.10. *Ogni formula A è equisoddisfacibile ad una formula in forma normale di Skolem con le stesse variabili libere.*

$\forall x \neg A \leftrightarrow \neg \exists x A$	De Morgan generalizzata
$\neg \forall x \neg A \leftrightarrow \exists x A$	De Morgan generalizzata
$\forall x A \leftrightarrow \neg \exists x \neg A$	De Morgan generalizzata
$\forall x A \rightarrow A[x/t]$	particolarizzazione (t sostituibile)
$A[x/t] \rightarrow \exists x A$	generalizzazione esistenziale (t sostituibile)
$\forall x A \leftrightarrow \forall y A[x/y]$	rinomina
$\exists x A \leftrightarrow \exists y A[x/y]$	rinomina
$\forall x(A \wedge B) \leftrightarrow \forall x A \wedge \forall x B$	distributività di $\forall$ su $\wedge$
$\exists x(A \vee B) \leftrightarrow \exists x A \vee \exists x B$	distributività di $\exists$ su $\vee$
$\forall x A \vee \forall x B \rightarrow \forall x(A \vee B)$	distributività parziale di $\forall$ su $\vee$
$\exists x(A \wedge B) \rightarrow \exists x A \wedge \exists x B$	distributività parziale di $\exists$ su $\wedge$
$\forall x(A \rightarrow B) \rightarrow \forall x A \rightarrow \forall x B$	distributività parziale di $\forall$ su $\rightarrow$
$\forall x(A \rightarrow B) \leftrightarrow A \rightarrow \forall x B$	(x non libera in A)
$\exists x(A \rightarrow B) \leftrightarrow A \rightarrow \exists x B$	(x non libera in A)
$\forall x(A \vee B) \leftrightarrow A \vee \forall x B$	(x non libera in A)
$\exists x(A \wedge B) \leftrightarrow A \wedge \exists x B$	(x non libera in A)
$\forall x(A \rightarrow B) \leftrightarrow \exists x A \rightarrow B$	(x non libera in B)
$\exists x(A \rightarrow B) \leftrightarrow \forall x A \rightarrow B$	(x non libera in B)
$\forall x \forall y A \leftrightarrow \forall y \forall x A$	scambio dei quantificatori
$\exists x \exists y A \leftrightarrow \exists y \exists x A$	scambio dei quantificatori

Tabella 7.1: Leggi logiche notevoli 2.

Dimostrazione. Data una formula A, la si mette in forma prenessa equivalente. Per eliminare dal prefisso i quantificatori esistenziali si procede in questo modo: se la formula è della forma $\forall x_1 \ldots \forall x_n \exists y B$, si introduce un nuovo simbolo funzionale F a n argomenti, e la formula si trasforma in $\forall x_1 \ldots \forall x_n B[y/F(x_1, \ldots, x_n)]$.

Se la forma prenessa è della forma $\exists y B$, si introduce una nuova costante c e si considera $B[y/c]$.

Eliminato un quantificatore esistenziale, si itera. $\square$

Esempio 7.2.11. L'enunciato in forma normale di Skolem che è equisoddisfacile a $\forall x \exists y R(x, y)$ è $\forall x R(x, F(x))$.

Quando si è trovata la forma universale equisoddisfacibile, la matrice si può sempre trasformare in forma normale congiuntiva. Ma dal punto di vista pratico non sempre conviene passare attraverso la forma prenessa; se la formula è una congiunzione $A_1 \wedge \ldots \wedge A_n$, conviene trasformare prima ogni A_i in forma normale di Skolem, e dopo farne la congiunzione. Che il risultato sia equisoddisfacibile segue dal fatto che se $A_1 \wedge \ldots \wedge A_n$ ha un modello $\mathcal{M}$, ciascuna delle forme di Skolem delle A_i è valida in un arricchimento, relativo a simboli funzionali diversi, della stessa $\mathcal{M}$.

Esempio 7.2.12. All'enunciato $\exists x \forall y P(x, y) \land \neg \exists z \forall x P(x, z)$ si può associare la forma $\forall y P(c, y) \land \forall z \neg P(F(z), z)$, equivalente a $\forall y \forall z (P(c, y) \land \neg P(F(z), z))$ invece di passare alla forma prenessa $\exists x \forall y \forall z \exists u (P(x, y) \land \neg P(u, z))$ e quindi a $\forall y \forall z (P(c, y) \land \neg P(G(y, z), z))$ dove il simbolo funzionale G è a due argomenti.

Esercizio 7.2.13. Trasformare in forma prenessa le seguenti formule:

$$\begin{gathered}
\forall x \exists y R(x, y) \land (\exists x P(x) \lor \exists x Q(x)) \\
\forall x \exists y R(x, y) \lor (\exists x P(x) \land \exists x Q(x)) \\
\forall x P(x) \lor \forall x (Q(x) \to \exists z R(x, z)) \\
\forall x (P(x) \lor \forall x Q(x)) \to \exists x (P(x) \lor Q(x)).
\end{gathered}$$

Esercizio 7.2.14. Trasformare in forma normale di Skolem le seguenti formule, anche in più di un modo:

$$\begin{gathered}
\exists x P(x) \lor \neg \exists x P(x) \\
\forall x P(x) \land \exists x \neg P(x) \\
\exists x P(x) \to \exists x Q(x) \\
\forall x \exists y (\forall z R(z, y) \lor \neg P(x, y)) \land \forall x \exists y \forall z \exists u (P(x, z) \to R(y, u)) \\
\forall x \exists y R(x, y) \to (\exists y \forall x R(x, y) \lor \forall y \exists x \neg R(x, y)) \\
\neg(\forall x \exists y R(x, y) \to (\exists y \forall x R(x, y) \lor \forall y \exists x \neg R(x, y))).
\end{gathered}$$

7.3 Quantificatori e dimostrazioni

Completata la presentazione del linguaggio dei predicati e delle principali leggi logiche, possiamo vedere come intervengono nelle dimostrazioni.

Le frasi matematiche presenti nelle premesse e conclusioni di una dimostrazione sono rappresentate da enunciati di linguaggi predicativi, mentre i passaggi intermedi di solito sono formule, formule algebriche[16] o loro combinazioni proposizionali, con variabili libere; si tratta di vedere come si fa a togliere e (ri)mettere i quantificatori.

Le regole per la deduzione naturale date nel capitolo 2 hanno questo scopo. Le ricordiamo precisando la condizione allora non menzionata che nell'eliminazione dell'universale e nell'introduzione dell'esistenziale il termine deve essere sostituibile nella formule.

[16]Una definizione generale di "formula algebrica" potrebbe essere "formula atomica con predicato $=$ o $<$".

(E∀) $$\frac{\forall x B[x]}{B[x/t]}$$

(I∃) $$\frac{B[x/t]}{\exists x B[x]}$$

(E∃) Se da $B[x/c]$ si deduce A, dove c è un nuovo simbolo che non occorre nella parte precedente della deduzione né in A, allora A si deduce da $\exists x B[x]$.

(I∀) Se $B[x]$ è dedotto da assunzioni nelle quali non occorre x priva della specificazione quantitativa data da un quantificatore $\forall x$ o $\exists x$, allora si può dedurre $\forall x B[x]$.

Prima di discutere (E∃) vediamo un esempio di esemplificazione esistenziale: dimostriamo che se un numero è divisibile per 4 allora è divisibile per 2. In questo caso l'elemento sconosciuto apparentemente si mantiene fino alla fine, nella conclusione, ed allora deve essere eliminato da questa con un'applicazione della generalizzazione esistenziale.

Da
$$\exists y(x = 4 \cdot y)$$
si passa a
$$\begin{aligned} x =& 4 \cdot c \\ x =& (2 \cdot 2) \cdot c \\ x =& 2 \cdot (2 \cdot c) \end{aligned}$$
Di qui si vede che x è divisibile per 2, ma non si può terminare con questa formula. Per generalizzazione esistenziale invece, considerata la formula $x = 2 \cdot (2 \cdot c)$ come formula del tipo $A[y/2 \cdot c]$, con $A[y]$ uguale a $x = 2 \cdot y$, si può allora dedurre
$$\exists y(x = 2 \cdot y),$$
per concludere
$$\exists y(x = 4 \cdot y) \to \exists y(x = 2 \cdot y)$$
e quindi
$$\forall x(\exists y(x = 4 \cdot y) \to \exists y(x = 2 \cdot y)).$$

Il problema logico della regola di esemplificazione esistenziale consiste nel fatto che $A[x/c]$ non è conseguenza logica di $\exists x A[x]$, e quindi nella successione di formule che costituiscono una dimostrazione che faccia uso della regola non è più vero che ogni formula è conseguenza logica delle precedenti (o un assioma del dominio in oggetto o una legge logica).

$A[x/c]$ non è conseguenza logica di $\exists x A[x]$ perché può succedere che in un'interpretazione $\exists x A[x]$ sia vero ma c non sia uno degli elementi che soddisfano A. Ad esempio $\exists y(0 < y)$ è vero in $\mathbb{N}$, ma se c denota 0 allora $0 < c$ è falso. È vero che noi affermiamo $A[x/c]$, ma questa si presenta come una nuova assunzione su c, non come una conseguenza di $\exists x A[x]$.

Si può tuttavia dimostrare che, nelle condizioni della regola applicata a $\exists x A[x]$, se B è una formula che non contiene c, e se B è conseguenza logica di $A[x/c]$ allora B è conseguenza logica di $\exists x A[x]$ — nonostante $A[x/c]$ non sia conseguenza logica di $\exists x A[x]$.

Una spiegazione, in breve, è la seguente. Posto che B segua da $A[x/c]$, dove c non occorre in B, si ha

$$A[x/c] \to B.$$

Ma di c non si sa nulla, a parte il fatto $A[x/c]$. Qualunque elemento denoti c va bene, e soddisfa $A[x/c] \to B$. Questo sembra potersi esprimere come

$$\forall c(A[x/c] \to B)$$

a parte l'errore sintattico di quantificare un simbolo che abbiamo chiamato costante. Ma se l'avessimo chiamato variabile, se avessimo usato una variabile w della nostra disponibilità illimitata di variabili, una che non occorresse altrimenti nella dimostrazione (in modo da non avere per w vincoli che non ci devono essere, salvo $A[x/w]$), avremmo concluso

$$A[x/w] \to B$$

con w non in B, quindi

$$\forall w(A[x/w] \to B)$$

e infine

$$\exists w A[x/w] \to B.$$

Ora ci limitiamo a far vedere che le applicazioni della regola di esemplificazione esistenziale possono essere sostituite da altri ragionamenti che non ne fanno uso.

Consideriamo l'esempio di

$$\exists y(x = 2 \cdot y) \to \neg\exists y(x + 1 = 2 \cdot y),$$

dove normalmente si assumerebbe $\exists y(x = 2 \cdot y)$ e si applicherebbe (E$\exists$).

Poiché y non è libera nel conseguente, si può scrivere in modo equivalente

$$\forall y(x = 2 \cdot y \to \neg\exists y(x + 1 = 2 \cdot y));$$

come nell'esempio dell'amico che rende felici, che può essere qualunque, anche ora si dice che y può essere qualunque, purché soddisfi poi l'antecedente $x = 2 \cdot y$.

Trattiamo la y di $x = 2 \cdot y$ è come una variabile universale. Data una y qualunque, occorre dimostrare che

$$x = 2 \cdot y \to \neg\exists y(x + 1 = 2 \cdot y),$$

ovvero

$$x = 2 \cdot y \to \forall y(x + 1 \neq 2 \cdot y).$$

Il quantificatore del conseguente può essere spostato nel prefisso, dopo aver eseguito un'opportuna rinomina, e la formula da dimostrare è equivalente a

$$\forall z(x = 2 \cdot y \to x + 1 \neq 2 \cdot z);$$

quindi possiamo provare a dimostrare

$$x = 2 \cdot y \to x + 1 \neq 2 \cdot z,$$

con tutte le variabili intese in senso universale.

Per assurdo, assumiamo la negazione del condizionale, quindi

$$x = 2 \cdot y \wedge x + 1 = 2 \cdot z,$$

e con alcuni calcoli arriviamo a una contraddizione.

Abbiamo quindi

$$x = 2 \cdot y \to x + 1 \neq 2 \cdot z,$$

e quantificando universalmente

$$\forall x \forall y \forall z(x = 2 \cdot y \to x + 1 \neq 2 \cdot z),$$

da cui con le leggi logiche pertinenti

$$\forall x(\exists y(x = 2 \cdot y) \to \forall z(x + 1 \neq 2 \cdot z)).$$

Si noti che di solito nel gergo matematico, dove non si usa indicare i quantificatori, attraverso un'interpretazione (corretta) dell'enunciato da dimostrare si imposta direttamente proprio

$$\text{se } x = 2 \cdot y, \text{ allora } x + 1 \neq 2 \cdot z.$$

La regola relativa all'eliminazione temporanea del quantificatore esistenziale afferma dunque che si può esemplificare un'affermazione esistenziale $\exists x A$ con $A[x/c]$ o $A[x/w]$ se per questa via si perviene a una conclusione che non contiene la costante o non contiene libera la variabile usata per l'esemplificazione.

Tuttavia le cautele e le restrizioni per questa regola non sono ancora finite.

Finché la costante o la variabile introdotte come esemplificazione di un quantificatore esistenziale non sono scomparse, l'argomento è incompleto, e non terminato, come in sospeso, per il riferimento a questo elemento sconosciuto. La costante o variabile può scomparire o per passaggi proposizionali (come sopra, una dimostrazione per assurdo di un altro enunciato, oppure per il taglio di un *modus ponens*), o per generalizzazione esistenziale.

Quello che bisogna assolutamente evitare è di quantificare universalmente una variabile che sia stata introdotta come esemplificazione di un quantificatore esistenziale (in questo l'uso di una costante ha ovvi vantaggi).

Un esempio di errore clamoroso dovuto a una simile disattenzione è la seguente dimostrazione di $\exists x \forall y(x < y)$ a partire da $\forall x \exists y(x < y)$.

Assunto $\forall x \exists y(x < y)$, per particolarizzazione si ha $\exists y(x < y)$; per esemplificazione esistenziale, sia y tale che $x < y$. Se ora dimenticandosi della natura esistenziale di y si affermasse $\forall y(x < y)$ si potrebbe concludere per generalizzazione esistenziale che $\exists x \forall y(x < y)$.

Ma questa conclusione non è conseguenza della premessa, come si vede dal fatto che la premessa è ad esempio vera negli interi, mentre la conclusione non lo è.

Anche la gestione della introduzione del quantificatore universale è più delicata di quanto finora abbiamo lasciato intendere. Si possono legittimamente (ri)quantificare universalmente le variabili libere che derivano per particolarizzazione da un quantificatore universale, ma non è questa tutta la storia. A volte sembra di lavorare con varibili libere che non derivano da una particolarizzazione, e che pure hanno un significato universale. La vera condizione è che le variabili non occorrano libere nelle assunzioni.

Ad esempio, se si parte assumendo $0 < x$ e con un argomento corretto, utilizzando le proprietà dei numeri reali, si conclude $\exists y(x = y^2)$, non si può affermare $\forall x \exists y(x = y^2)$ — c'è una condizione restrittiva su x stabilita dalla premessa. In realtà l'argomento che porta da $0 < x$ a $\exists y(x = y^2)$ stabilisce $0 < x \to \exists y(x = y^2)$ per x qualunque, senza alcuna premessa (salvo le proprietà dei numeri reali espresse da enunciati, senza variabili libere). Quindi x non è libera nelle premesse della derivazione di quest'ultima formula, che non ci sono, e si può correttamente quantificarla in $\forall x(0 < x \to \exists y(x = y^2))$.

Infine esiste un problema con la generalizzazione universale all'interno della esemplificazione esistenziale.

Un esempio di errore dovuto a cattiva gestione della quantificazione universale è il seguente. Da $\forall x \exists y(x < y)$ per particolarizzazione si ha $\exists y(x < y)$ e per esemplificazione, sia c tale che $x < c$. Se ora si quantifica universalmente x si ottiene $\forall x(x < c)$ e per generalizzazione esistenziale $\exists y \forall x(x < y)$. La conclusione, che afferma l'esistenza di un massimo per $<$, è palesemente falsa nei naturali, dove invece la premessa è vera. Ma la premessa è un enunciato, e sembrerebbe quindi che non vi fossero variabili libere nelle premesse e tutte le regole fossero applicate correttamente. La spiegazione sta nel fatto che quanto si dice "sia c tale che $x < c$" inizia, come abbiamo detto sopra, un argomento particolare, una sorta di dimostrazione a parte che non si considera conclusa finché tale c non sparisce legittimamente. In questa dimostrazione subordinata, $x < c$ è una premessa, e x è libera in $x < c$, e non è lecito perciò quantificare universalmente la x.

Esercizio 7.3.1. Dedurre nel calcolo della deduzione naturale la generalizzazione esistenziale dalla particolarizzazione universale e da De Morgan generalizzata.

Esercizio 7.3.2. Dedurre nel calcolo della deduzione naturale le leggi di rinomina dalla legge di particolarizzazione e dalle leggi relative a condizionale e quantificatori.

Esercizio 7.3.3. Dedurre nel calcolo della deduzione naturale le une dalle altre le leggi riguardanti $\forall$ e $\vee$, $\exists$ e $\wedge$, $\forall$ e $\rightarrow$.

Capitolo 8

Calcoli logici

8.1 Alberi di refutazione

La tecnica degli alberi di refutazione si può estendere agli enunciati dei linguaggi predicativi aggiungendo le seguenti regole:

- Se A è $\exists xB$, si introduce una nuova costante c e alla fine di ogni ramo non chiuso passante per A si appende alla foglia il successore $B[x/c]$, come nello schema

$$\begin{array}{c} [\exists xB] \\ \\ \vdots \\ \downarrow \\ B[x/c] \end{array}$$

- Se A è $\neg\forall xB$, si introduce una nuova costante c e alla fine di ogni ramo non chiuso passante per A si appende alla foglia il successore $\neg B[x/c]$, come nello schema

$$\begin{array}{c} [\neg\forall xB] \\ \\ \vdots \\ \downarrow \\ \neg B[x/c] \end{array}$$

- Se A Se A è $\forall xB$, allora alla fine di ogni ramo non chiuso passante per A, per tutti i termini chiusi $t_1, \ldots, t_n$ che occorrono in qualche enunciato del ramo, e tali che $B[x/t_i]$ non occorre già nel ramo, si appendono alla foglia $n+1$ nodi in serie, prima $B[x/t_1], \ldots, B[x/t_n]$ e poi ancora $\forall xB$, come nello schema

$$
\begin{array}{c}
[\forall xB] \\
\\
\vdots \\
\downarrow \\
B[x/t_1] \\
\downarrow \\
\vdots \\
\downarrow \\
B[x/t_n] \\
\downarrow \\
\forall xB
\end{array}
$$

- Se A Se A è $\neg\exists xB$, allora alla fine di ogni ramo non chiuso passante per A, per tutti i termini chiusi $t_1, \ldots, t_n$ che occorrono in qualche enunciato del ramo, e tali che $B[x/t_i]$ non occorre già nel ramo, si appendono alla foglia $n+1$ nodi in serie $\neg B[x/t_1], \ldots, \neg B[x/t_n]$, e poi ancora $\neg\exists xB$, come nello schema

$$
\begin{array}{c}
[\neg\exists xB] \\
\downarrow \\
\neg B[x/t_1] \\
\downarrow \\
\vdots \\
\downarrow \\
\neg B[x/t_n] \\
\downarrow \\
\neg\exists xB
\end{array}
$$

Se l'albero è inizializzato con un enunciato, tutti i nodi dell'albero sono etichettati con enunciati, di un linguaggio possibilmente arricchito con nuove costanti. Il ruolo dei letterali è ora svolto dagli enunciati atomici e dalle negazioni degli enunciati atomici.

Il senso delle due ultime regole è il seguente; per la regola relativa al quantificatore universale (per la negazione dell'esistenziale valgono le stesse considerazioni) si vorrebbero sostituire a x in B tutti i termini chiusi; ma questi sono in generale infiniti, e neppure ben determinati, per il fatto che successive applicazioni delle altre regole ad altri nodi possono introdurre nuove costanti; allora si incominciano a sostituire i termini esplicitamente esistenti, ma si riscrive l'enunciato $\forall xB$ in modo che quando eventualmente (se il ramo non si è nel frattempo chiuso) si torna a considerare l'enunciato, se nel frattempo si sono creati nuovi termini chiusi anche i nuovi vengano sostituiti.

Se in una prima applicazione delle ultime due regole non esistono termini chiusi negli enunciati dell'albero, si introduce una nuova costante c e si sostituisce quella.

Praticamente, non c'è bisogno di riscrivere $\forall xB$, basta non marcarlo come già considerato e ricordarsi di tornare periodicamente a visitarlo. E quando tutti gli altri enunciati siano stati considerati e non ci siano altri termini da sostituire, lo si marca definitivamente per terminare.

Esempio 8.1.1. Un albero chiuso:

$$
\begin{tikzcd}[column sep=small]
 & \neg(\forall x(P(x) \to Q(x)) \to (\forall xP(x) \to \forall xQ(x)))_1 \arrow[d] & \\
 & \forall x(P(x) \to Q(x))_5 \arrow[d] & \\
 & \neg(\forall xP(x) \to \forall xQ(x))_2 \arrow[d] & \\
 & \forall xP(x)_4 \arrow[d] & \\
 & \neg\forall xQ(x)_3 \arrow[d] & \\
 & \neg Q(c) \arrow[d] & \\
 & P(c) \arrow[d] & \\
 & P(c) \to Q(c) \arrow[dl] \arrow[dr] & \\
 \underset{\text{chiuso}}{\neg P(c)} & & \underset{\text{chiuso}}{Q(c)}
\end{tikzcd}
$$

Non vale più la proprietà di terminazione, come mostra l'albero per l'enunciato $\forall x \exists y R(x, y)$

$$\begin{array}{c} \forall x \exists y R(x,y) \\ \downarrow \\ \exists y R(c,y) \\ \downarrow \\ R(c,c_1) \\ \downarrow \\ \exists y R(c_1,y) \\ \downarrow \\ R(c_1,c_2) \\ \downarrow \\ \exists y R(c_2,y) \\ \downarrow \\ \vdots \end{array}$$

Valgono però le proprietà fondamentali di correttezza e completezza.

Teorema 8.1.2 (Correttezza)**.** *Se l'albero di refutazione con radice A si chiude, allora A è insoddisfacibile.* $\Box$

Teorema 8.1.3 (Completezza)**.** *Se A è insoddisfacibile, l'albero con radice A si chiude.*

La dimostrazione segue dal

Lemma 8.1.4. *Se l'albero di refutazione con radice A non si chiude, allora per ogni ramo non chiuso, finito e terminato, o infinito, esiste un modello di A.*

Dimostrazione. In ogni caso si definisce un'interpretazione $\mathcal{M}$ nel seguente modo. L'universo M è l'insieme dei termini chiusi che occorrono in qualche enunciato del ramo (anche sottotermini di termini). L'interpretazione è definita solo per i simboli che effettivamente compaiono in qualche enunciato del ramo. Per ogni costante c, dell'alfabeto originario o introdotta nel corso del processo, si pone $c^M = c$; per ogni simbolo funzionale F a n argomenti e ogni $t_1, \ldots, t_n \in M$, si pone $F^M(t_1, \ldots, t_n) = F(t_1, \ldots, t_n)$ se $F(t_1, \ldots, t_n)$ appartiene a M, altrimenti un termine qualunque. Si ha allora che $t^M = t$ per ogni $t \in M$.

Infine per ogni simbolo predicativo P a n argomenti e ogni $t_1, \ldots, t_n \in M$ si pone per definizione

$$\langle t_1, \ldots, t_n \rangle \in P^M \text{ se e solo se } P(t_1, \ldots, t_n) \text{ è un nodo del ramo.}$$

Con questa definizione, si verifica facilmente, per induzione sull'altezza degli enunciati, che per ogni enunciato A che occorre nel ramo si ha $\mathcal{M} \models A$. $\square$

Esempio 8.1.5. L'albero

$$\begin{array}{c} P(c) \land (\exists x P(x) \to \exists x Q(x)) \\ \downarrow \\ P(c) \\ \downarrow \\ \exists x P(x) \to \exists x Q(x) \\ \swarrow \qquad \searrow \\ \begin{array}{cc} \neg\exists x P(x) & \exists x Q(x) \\ \downarrow & \downarrow \\ \neg P(c) & Q(d) \end{array} \end{array}$$

mostra che l'enunciato $P(c) \land (\exists x P(x) \to \exists x Q(x))$ è soddisfacibile con un modello (dato dal ramo di destra, quello di sinistra è chiuso) costituito da $M = \{c, d\}$, $c^M = c$, $d^M = d$ e $P^M = \{c\}$ e $Q^M = \{d\}$.

Gli alberi di refutazione si possono usare per stabilire la validità logica di formule aperte considerando la loro chiusura universale. Se, data A con la sola variabile libera x, per esempio, si considera $\forall x A$ e per verificarne la verità logica si inizializza un albero con $\neg\forall x A$, la prima regola introduce una nuova costante c e l'enunciato $\neg A[x/c]$. Se invece che da $\forall x A$ si parte da $A[x/c]$, la chiusura dell'albero stabilisce la verità logica di $A[x/c]$, dove però c, essendo nuova e quindi non ristretta da A o da sue sottoformule a denotare uno specifico elemento, svolge il ruolo di elemento arbitrario, come le variabili. Ma se si parte da $A[x/c]$ e l'albero non si chiude, ciò non significa che A è soddisfacibile, perché $A[x/c]$ ha un modello con un particolare valore per c; significa che $\exists x A$ lo è. Di fatto si potrebbe definire il metodo in modo da permettere anche formule aperte, risulta solo un po' più faticosa la dimostrazione della correttezza e completezza.

Esercizio 8.1.6. Verificare con gli alberi di refutazione tutte le leggi logiche finora incontrate.

Esercizio 8.1.7. Utilizzare gli alberi di refutazione per spiegare che se una formula ha la struttura proposizionale di una tautologia allora è logicamente valida.

Esercizio 8.1.8. Trovare un controesempio a $\exists x P(x) \land \exists x Q(x) \to \exists x (P(x) \land Q(x))$.

Esercizio 8.1.9. Trovare un controesempio a $\forall x (P(x) \lor Q(x)) \to \forall x P(x) \lor \forall x Q(x)$.

Esercizio 8.1.10. Trovare un controesempio a $\forall x P(x) \to \forall x Q(x) \models \forall x(P(x) \to Q(x))$.

Esercizio 8.1.11. Verificare se i seguenti enunciati sono soddisfacibili, e in caso positivo descrivere i loro modelli, e quali sono infiniti:

$$\forall x \exists y R(x,y) \to \exists y \forall x R(x,y)$$
$$\forall x R(c,x) \to \forall x \exists y R(x,y)$$
$$\neg \forall x P(x) \wedge \forall x(Q(x) \to P(x)) \wedge \forall x Q(x)$$
$$\forall x R(c,x) \wedge \forall x \neg R(F(x),c) \wedge \forall x R(x,F(x))$$

Esercizio 8.1.12. Dimostrare che se l'enunciato della radice contiene solo simboli relazionali a un posto e nessun simbolo funzionale (linguaggio monadico) l'albero di refutazione è sempre finito.

Suggerimento. Dimostrare prima con le leggi del paragrafo 8.5 che ogni enunciato del genere è logicamente equivalente ad uno in cui non ci sono quantificatori nidificati, cioè nessun quantificatore cade dentro al raggio d'azione di un altro.

8.2 Calcolo della risoluzione

8.2.1 Interpretazioni di Skolem-Herbrand

Le interpretazioni definite nella dimostrazione del lemma 8.1.4, per ogni ramo non chiuso, sono particolari: l'universo è un insieme di termini chiusi; le funzioni sono definite in modo che risulti che ogni termine denota se stesso, $t^M = t$ e le relazioni sono definite scegliendo un insieme di enunciati atomici come veri. Interpretazioni con queste caratteristiche si chiamano interpretazioni *di Skolem-Herbrand*.

Avvertenza 8.2.1. Prima di pensare a chissà quali diavolerie, si consideri che le interpretazioni in cui i termini chiusi denotano se stessi sono le più naturali. Nelle prime esperienze con i numeri, le cifre $1, 2, \ldots$ non denotano numeri, ma *sono* i numeri, ovvero questi si identificano con le cifre. Solo a livelli sofisticati si introduce una notazione diversa per i numeri denotati dalle cifre o dai numerali; ad esempio l'insieme vuoto $\emptyset$ è il numero zero secondo la definizione insiemistica, e 0 la cifra del linguaggio aritmetico che lo denota, oppure in un'interpretazione $\mathcal{M}$ dell'aritmetica l'elemento 0^M di M è lo zero della struttura, e 0 sempre la costante del linguaggio che lo denota.

La differenza è che nella trattazione intuitiva tendiamo a dire che ad esempio $1 + 1$ denota 2, mentre in un'interpretazione di Skolem-Herbrand

$1 + 1$ denota se stesso, e *si riduce* alla forma normale 2, ovvero, $1 + 1$ e 2 stanno nella stessa classe di equivalenza rispetto a $=$.

Dal lemma 8.1.4 si può perciò dedurre

Corollario 8.2.2. *Se un enunciato A è soddisfacibile, A ha un modello di Skolem-Herbrand*

osservando che se A è soddisfacibile, l'albero con radice A non si chiude, e A ha modelli costruiti come nella dimostrazione del lemma.

Un'interpretazione di Skolem-Herbrand si può vedere come interpretazione proposizionale degli enunciati privi di quantificatori: la scelta degli enunciati atomici veri che individua l'interpretazione di Skolem-Herbrand, e che si chiama anche *base* di Herbrand (che nel caso dell'albero di refutazione sono quelli nei nodi di un ramo non chiuso, ma in altri contesti potrebbero essere scelti in modo diverso) si può pensare come l'attribuzione del valore 1 ad alcuni enunciati atomici (agli altri il valore 0); gli enunciati privi di quantificatori sono ottenuti da quelli atomici con i connettivi e il loro essere veri o falsi nell'interpretazione di Skolem-Herbrand coincide con il loro avere 1 o 0 come valore calcolato con le funzioni di verità.

Un insieme $\mathcal{S}$ di enunciati privi di quantificatori si dice *soddisfacibile in senso proposizionale* se esiste un'assegnazione v di valori $0, 1$ agli enunciati atomici occorrenti come sottoenunciati degli enunciati di $\mathcal{S}$ tale che $v^*(A) = 1$ per ogni $A \in \mathcal{S}$.

Per un enunciato universale $\forall x_1 \ldots \forall x_n A$, sia $\mathcal{S}_A$ l'insieme degli enunciati privi di quantificatori che si ottengono sostituendo i termini chiusi in tutti i modi possibili nella matrice A.

Teorema 8.2.3 (Teorema di Skolem-Herbrand)**.** *Un enunciato universale $\forall x_1 \ldots \forall x_n A$ è insoddisfacibile se e solo se esiste un sottinsieme finito di $\mathcal{S}_A$ che è insoddisfacibile in senso proposizionale.*

Dimostrazione. Se $\forall x_1 \ldots \forall x_n A$ è insoddisfacibile, l'albero che ha la radice $\forall x_1 \ldots \forall x_n A$ si chiude. Nello sviluppo dell'albero si applicherà prima che si chiuda un numero finito di volte la regola per il quantificatore universale, che introduce nell'albero enunciati $A[x_1/t_1^i, \ldots, x_n/t_n^i]$ a cui poi si applicano solo regole relative a connettivi.

Se si imposta un albero con un ramo iniziale in cui tutti questi enunciati $A[x_1/t_1^i, \ldots, x_n/t_n^i]$ sono introdotti in serie, nello sviluppo successivo di quest'albero si applicano le stesse mosse che si sono applicate nel precedente, e che portano alla sua chiusura.

Viceversa se c'è un insieme finito di enunciati $A[x_1/t_1^i, \ldots, x_n/t_n^i] \in \mathcal{S}_A$ tale che l'albero che ha come radice la loro congiunzione si chiude, perché l'insieme è insoddisfacibile, allora per l'albero con radice $\forall x_1 \ldots \forall x_n A$ nel suo sviluppo si può iterare la regola del quantificatore universale, prima di quelle relative ai connettivi (i nuovi termini vengono solo dalla sostituzione di termini in altri termini, perché non si applica né la regola del quantificatore esistenziale né quella della negazione dell'universale), finché prima o poi si introdurranno gli $A[x_1/t_1^i, \ldots, x_n/t_n^i]$ che con il loro sviluppo seguente portano alla chiusura dell'albero. □

Gli elementi di $\mathcal{S}_A$ si chiamano anche *esemplificazioni di base* (ingl. *ground instances*) di A, o della chiusura universale $\forall A$ di A.

Esempio 8.2.4. Se $\forall x A$ è $\forall x(R(c,x) \wedge \neg R(x, F(x)))$, allora il sottinsieme di $\mathcal{S}_A$ insoddisfacibile in senso proposizionale è

$$\{R(c,c) \wedge \neg R(c, F(c)), R(c, F(c)) \wedge \neg R(F(c), F(F(c)))\}$$

che si ottiene con due sostituzioni ed ha la struttura proposizionale $\{P \wedge \neg Q, Q \wedge R\}$.

Esercizio 8.2.5. Applicando la forma normale di Skolem e il teorema di Skolem-Herbrand, mostrare che la congiunzione dei due enunciati

$$\forall x \forall y \forall z(R(x,y) \wedge R(y,z) \to R(x,z))$$
$$\exists y(R(c,y) \wedge R(y,c) \wedge \neg R(y,y))$$

è una contraddizione.

Esercizio 8.2.6. Scrivere in forma normale di Skolem l'enunciato

$$\exists y \forall x(R(y,x) \wedge \exists z(R(x,z) \to \neg R(x,x)))$$

e verificare con gli alberi di refutazione che è insoddisfacibile trovando le sostituzioni di termini chiusi che forniscono l'insieme proposizionalmente insoddisfacibile di esemplificazioni di base.

8.3 Unificazione

Per ottenere un insieme finito di esemplificazioni di base che sia proposizionalmente insoddisfacibile occorre talvolta fare diverse sostituzioni, e solo l'effetto combinato di più di una fornisce l'insieme cercato.

Esempio 8.3.1. Da $\forall x \exists y \forall z(P(c,x) \wedge \neg P(z,y)))$ si ottiene la forma normale di Skolem $\forall x \forall z(P(c,x) \wedge \neg P(z, F(x)))$. La prima sostituzione di c a x e c a z fornisce

$$P(c,c) \wedge \neg P(c, F(c))$$

una seconda di c a x e $F(c)$ a z fornisce

$$P(c,c) \wedge \neg P(F(c), F(c))$$

e una terza di $F(c)$ a x e c a z

$$P(c, F(c)) \wedge \neg P(c, F(F(c)))$$

da cui si vede la contraddizione tra $\neg P(c, F(c))$ della prima esemplificazione e $P(c, F(c))$ della terza.

Se invece eseguiamo le seguenti trasformazioni equivalenti della forma normale di Skolem:

$$\begin{gathered}
\forall x \forall z (P(c,x) \wedge \neg P(z, F(x))) \\
\forall x \forall z P(c,x) \wedge \forall x \forall z \neg P(z, F(x)) \\
\forall x P(c,x) \wedge \forall y \forall z \neg P(z, F(y)) \\
\forall x \forall y \forall z (P(c,x) \wedge \neg P(z, F(y)))
\end{gathered}$$

arriviamo alla matrice

$$P(c,x) \wedge \neg P(z, F(y))$$

dove è sufficiente l'unica sostituzione di $F(c)$ a x, c a y e c a z.

Con i passaggi eseguiti nell'esempio, è sempre possibile avere una matrice, che d'ora in poi chiameremo insieme di causole, che sia *a variabili disgiunte*, cioè tale che nessuna variabile occorra in due clausole.

Per non dover applicare sistematicamente tutte le sostituzioni possibili, occorre individuare quelle utili, che, nella prospettiva di applicare il metodo di risoluzione per stabilire l'insoddisfacibilità, sono quelle che fanno comparire letterali complementari in clausole diverse. Si tratta di prevedere quali sostituzioni, se ne esistono, renderanno uguali due formule atomiche. Questa previsione e l'individuazione della sostituzione è possibile con un algoritmo, che costruisce la sostituzione per approssimazioni successive, componendo successive sostituzioni, e quindi lavora con termini che contengono variabili.

Esempio 8.3.2. Per la forma normale congiuntiva

$$(P(x) \vee R(x)) \wedge \neg P(F(y)) \wedge \neg R(F(G(c)))$$

la sostituzione che produce un enunciato insoddisfacibile è quella che a x sostituisce $F(G(c))$ e a y sostituisce $G(c)$. Ma questa la si può riconoscere considerando prima la sostituzione di $F(y)$ a x e quindi la sostituzione di $G(c)$ a y.

Una sostituzione simultanea del termine t_1 alla variabile x_1, ..., e di t_n a x_x si indicherà con la notazione $\{x_1/t_1, \ldots, x_n/t_n\}$ e sostituzioni del genere con lettere greche. L'applicazione di una sostituzione σ a un'espressione W si indicherà con $W\sigma$; la composizione di due sostituzioni σ e θ con $\sigma\circ\theta$, che è la sostituzione tale che $W(\sigma\circ\theta) = (W\sigma)\theta$.

Un insieme finito $W = \{E_1, \ldots, E_n\}$ di espressioni (che per noi saranno sempre formule atomiche) si dice *unificabile* se esiste una sostituzione σ tale che $W\sigma = \{E_1\sigma, \ldots, E_n\sigma\}$ ha un solo elemento, cioè σ rende uguali le espressioni di W a cui si applica. σ si chiama *unificatore* di W.

8.3.1 Algoritmo di unificazione

Dato un insieme W di espressioni, si supponga di avere generato allo stadio k una sostituzione σ_k e un insieme di espressioni W_k tali che $W_k = W\sigma_k$. Allo stadio iniziale si può supporre la sostituzione identica.

Se W_k ha un solo elemento, W è unificabile e σ_k un unificatore; altrimenti, si esaminino le espressioni di W_k da sinistra a destra finché si trova il primo punto di disaccordo, dove non tutte le liste sono uguali, e si definisca l'*insieme di disaccordo* D_k come insieme dei termini, uno per ciascuna espressione di W_k, che iniziano nel punto di disaccordo. Se D_k contiene due termini, uno dei quali è una variabile v_k e l'altro è un termine t_k che non contiene v_k, si ponga $\sigma_{k+1} = \sigma_k\circ\{v_k/t_k\}$ e $W_{k+1} = W_k\{v_k/t_k\}$ (da cui segue $W_{k+1} = W\sigma_{k+1}$).

Se in D_k non esiste una tale coppia di termini, si esce dicendo che W non è unificabile. Se in D_k ci sono diverse variabili $v, \ldots, u$ che non occorre in un termine t_k si può anche porre $\sigma_{k+1} = \{v/t_k, \ldots, u/t_k\}$.

Quindi si itera.

L'algoritmo termina sempre perché a ogni stadio in cui si itera una variabile scompare, essendo sostituita ovunque da termini che non la contengono, e le variabili iniziali sono finite.

Esempio 8.3.3. Se $W = \{P(x, F(x), v), P(y, z, F(y)), P(u, F(G(w)), F(G(w)))\}$, o per maggior chiarezza

$$\begin{gathered} P(x, F(x), v) \\ (y, z, F(y)) \\ P(u, F(G(w)), F(G(w))) \end{gathered}$$

posto $W_0 = W$ si ha che $D_0 = \{x, y, u\}$ e si può porre $\sigma_1 = \{x/u, y/u\}$

ottenendo come W_1

$$P(u, F(u), v)$$
$$P(u, z, F(u))$$
$$P(u, F(G(w)), F(G(w))).$$

D_1 è ora $\{F(u), z, F(G(w))\}$ e si può porre $\sigma_2 = \sigma_1 \circ \{z/F(u)\}$ ottenendo per W_2

$$P(u, F(u), v)$$
$$P(u, F(u), F(u))$$
$$P(u, F(G(w)), F(G(w))).$$

$D_2 = \{u, G(w)\}$ e $\sigma_3 = \sigma_2 \circ \{u/G(w)\}$, e W_3 è

$$P(G(w), F(G(w)), v)$$
$$P(G(w), F(G(w)), F(G(w)))$$
$$P(G(w), F(G(w)), F(G(w))).$$

$D_3 = \{v, F(G(w))\}$, $\sigma_4 = \sigma_3 \circ \{v/F(G(w))\}$ e si ha infine

$$W_4 = \{P(G(w), F(G(w)), F(G(w)))\}.$$

Il risultato σ_4 delle successive composizioni si scrive

$$\sigma_4 = \{x/G(w), y/G(w), z/F(G(w)), u/G(w), v/F(G(w))\}.$$

Ovviamente se l'algoritmo fornisce un unificatore, W è unificabile. Viceversa, se W è unificabile, l'algoritmo fornisce un unificatore, anzi uno con proprietà aggiuntiva.

Nell'esempio, dato l'unificatore fornito dall'algoritmo, è subito visto che anche $\{x/G(s), y/G(s), z/F(G(s)), u/G(s), v/F(G(s))\}$, dove s è una nuova variabile, o $\{x/G(c), y/G(c), z/F(G(c)), u/G(c), v/F(G(c))\}$, o più in generale

$$\{x/G(t), y/G(t), z/F(G(t)), u/G(t), v/F(G(t))\}$$

per un termine qualunque t, cioè $\sigma_4 \circ \{w/t\}$, sono unificatori. Ma vale anche il viceversa.

Un unificatore σ di un insieme W si dice *generale* se per ogni altro unificatore θ di W esiste una sostituzione λ tale che $\theta = \sigma \circ \lambda$.

Un unificatore generale non è mai unico. In particolare si può eseguire una rinomina, sostituzione biunivoca delle variabili che compaiono nei termini dell'unificatore con nuove variabili, e si ottiene ancora un unificatore generale. Due unificatori generali differiscono solo per una rinomina.

L'unificatore fornito dall'algoritmo di unificazione è un unificatore generale.

Esempio 8.3.4. $W = \{P(x, F(x), v), P(c, z, F(y)), P(u, F(G(w)), F(G(w)))\}$ non è unificabile. I primi passi analoghi al precedente esempio portano a

$$\begin{gathered} P(c, F(c), v) \\ P(c, F(c), F(c)) \\ P(c, F(G(w)), F(G(w))) \end{gathered}$$

con le sostituzioni $\{x/c, u/c\} \circ \{z/F(c)\}$, ma ora $D_2 = \{c, G(w)\}$ e non sono soddisfatte le condizioni per proseguire.

Esempio 8.3.5. $W = \{P(x, z, G(x)), P(y, F(x), y)\}$ non è unificabile perché al terzo passo l'insieme di disaccordo è $\{G(x), x\}$ oppure $\{G(y), y\}$; l'esito negativo non dipende dal fatto che le due espressioni non sono a variabili disgiunte; lo stesso succede se la seconda è $P(y, F(u), y)$.

Esercizio 8.3.6. Applicare l'algoritmo di unificazione per decidere se i seguenti insiemi di espressioni sono unificabili (e trovare un unificatore generale) oppure no:

$$\begin{gathered} \{P(x, F(y), z), P(G(c), F(w), u), P(v, F(d), e)\} \\ \{P(F(x, y), w), P(F(G(v), c), H(v)), P(F(G(v), c), H(d))\} \\ \{P(x, F(x)), P(y, y)\} \\ \{P(H(y), c, z), P(H(F(w)), c, w), P(H(F(c)), c, d)\} \\ \{P(H(y), c, z), P(H(F(w)), c, w), P(H(F(c)), c, u)\} \\ \{P(F(x), y), P(y, y), P(y, G(z))\} \\ \{P(y, z, F(y)), P(x, F(x), v), P(u, F(G(w)), F(G(w))) \\ \{P(y, z, F(y)), P(x, F(x), v), P(u, F(G(w)), F(G(w))) \\ \{P(c, x, F(G(y))), P(z, F(y), F(y)), P(u, F(y), v)\}. \end{gathered}$$

8.3.2 Risoluzione con variabili

Dato un insieme di clausole con variabili, che si può sempre pensare sia la matrice di una forma normale di Skolem, si dirà che l'insieme di clausole è insoddisfacibile se è insoddisfacibile l'enunciato che ha quella forma normale di Skolem. L'insoddisfacibilità si può stabilire derivando la clausola vuota con una generalizzazione della regola di risoluzione.

Due letterali si chiamano *simili* se iniziano con lo stesso simbolo predicativo o iniziano entrambi con la negazione dello stesso simbolo predicativo; si chiamano *quasi-complementari* se uno inizia con un simbolo predicativo e l'altro con la negazione dello stesso simbolo, e si indicano con $L(t_1, \ldots, t_n)$ e $L^c(s_1, \ldots, s_n)$.

Definizione 8.3.7. Date due clausole a variabili disgiunte che contengono l'una letterali quasi-complementari di letterali dell'altra,

$$C_1 \cup \{L(t_{11}, \ldots, t_{1n}), \ldots, L(t_{h1}, \ldots, t_{hn})\}$$

e

$$C_2 \cup \{L^c(s_{11}, \ldots, s_{1n}), \ldots, L^c(s_{k1}, \ldots, s_{kn})\}$$

dove $\{L(t_{11}, \ldots, t_{1n}), \ldots, L(t_{h1}, \ldots, t_{hn})\}$ è un insieme di letterali simili in una delle due clausole e $\{L^c(s_{11}, \ldots, s_{1n}), \ldots, L^c(s_{k1}, \ldots, s_{kn})\}$ è un insieme di letterali quasi-complementari dei primi nella seconda, allora se σ è un unificatore generale di

$$\{L(t_{11}, \ldots, t_{1n}), \ldots, L(t_{h1}, \ldots, t_{hn}), L(s_{11}, \ldots, s_{1n}), \ldots, L(s_{k1}, \ldots, s_{kn})\}$$

la risolvente delle due clausole è $(C_1\sigma \cup C_2\sigma)\lambda$, dove λ è una rinomina tale che la clausola risolvente abbia variabili disgiunte dalle genitrici (e, quando inserita in una derivazione da un insieme $\mathcal{S}$, da tutte le clausole di $\mathcal{S}$ e da quelle già ottenute per risoluzione).

Esempio 8.3.8. La necessità, e non solo l'opportunità, di avere variabili disgiunte è illustrata dall'insieme di clausole $\{P(x), \neg P(F(x))\}$ da cui non si potrebbe derivare la clausola vuota con la risoluzione perché $P(x)$ e $P(F(x))$ non sono unificabili; invece tale insieme corrisponde all'enunciato $\forall x \exists y (P(x) \land \neg P(y))$ che è insoddisfacibile.

D'altra parte l'insieme $\{P(x), \neg P(F(y))\}$ che si può sostituire a quello dato è la matrice di $\forall x \forall y (P(x) \land \neg P(F(y)))$ che si può pensare proveniene da $\forall x P(x) \land \forall y \exists z \neg P(z) \equiv \forall x P(x) \land \exists z \neg P(z) \equiv \forall x P(x) \land \neg \forall z P(z)$.

Esempio 8.3.9. La necessità che anche le nuove clausole risolventi siano a variabili disgiunte da quelle già disponibili è illustrata dall'esempio dell'insieme di clausole $\{P(x) \lor P(F(x)), \neg P(y) \lor Q(y), \neg Q(z)\}$, associate all'enunciato insoddisfacibile $\forall x \exists y (P(x) \lor P(y)) \land \forall x (P(x) \to Q(x)) \land \neg \exists z Q(z)$.

Supponiamo di non eseguire le rinomine nelle clausole risolventi. Dalla prima e seconda, con l'unificatore $\{x/y\}$, si ottiene la risolvente $P(F(y)) \lor Q(y)$. Se ora si risolvesse questa con la terza rispetto al letterale $Q(y)$ con l'unificatore $\{z/y\}$ si avrebbe $P(F(y))$ che non potrebbe essere risolto con la seconda per la non unificabilità di $P(y)$ e $P(F(y))$. Se si usasse invece l'unificatore $\{y/z\}$, si otterrebbe $P(F(z))$ che risolto con la seconda, con l'unificatore $\{y/F(z)\}$ darebbe $Q(F(z))$ e si sarebbe di nuovo bloccati.

Se si partisse risolvendo la seconda e la terza, con l'unificatore $\{y/z\}$ si otterrebbe $\neg P(z)$ che con la prima e l'unificatore $\{x/z\}$ darebbe $P(F(z))$; questa con la seconda e l'unificatore $\{y/F(z)\}$ darebbe $Q(F(z))$ non unificabile con la terza clausola.

Se invece in questa derivazione, da $\neg P(z)$ si risolvesse con la prima con l'unificatore $\{z/x\}$ si otterrebbe $P(F(x))$, e questa con la seconda e l'unificatore $\{y/F(x)\}$ darebbe $Q(F(x))$ risolvibile infine con la terza. Oppure $P(F(x))$ si risolve con la clausola precedentemente ottenuta $\neg P(z)$.

Ma quando si voglia meccanizzare il procedimento, l'algoritmo di unificazione è fissato, e in esso un ordine in cui considerare le variabili e scegliere ad esempio tra $\{x/z\}$ e $\{z/x\}$, non si può modificarlo secondo astuzia o convenienza. Alcune euristiche si possono inserire nel programma, ad esempio tra due unificatori $\{x/z\}$ e $\{z/x\}$ preferire quello in cui la variabile da sostituire ha meno occorrenze, ma una volta fissate, queste vengono applicate. Programmare l'applicazione di tutti i possibili unificatori peraltro non sembra la soluzione più ragionevole.

Se ora torniamo alla prima risoluzione ed eseguiamo la rinomina $\{y/v\}$, possiamo risolvere la clausola ottenuta con la terza (usando indifferentemente sia $\{v/z\}$ sia $\{z/v\}$) e procedere fino alla clausola vuota.

Schematicamente, se numeriamo le tre clausole

$$\begin{array}{lc}
1 & P(x) \vee P(F(x)) \\
2 & \neg P(y) \vee Q(y) \\
3 & \neg Q(z)
\end{array}$$

possiamo rappresentare la derivazione con

$$\begin{array}{rl}
1 & \quad 2 \\
\downarrow & \swarrow_{\{x/y\}} \\
P(F(y)) \vee Q(y) & \\
P(F(v)) \vee Q(v) & \quad 3 \\
\downarrow & \swarrow_{\{z/v\}} \\
P(F(v)) & \\
P(F(u)) & \quad 2 \\
\downarrow & \swarrow_{\{y/F(u)\}} \\
Q(F(u)) & \\
Q(F(w)) & \quad 3 \\
\downarrow & \swarrow_{\{z/F(w)\}} \\
\Box &
\end{array}$$

La condizione di avere sempre tutte le clausole a variabili disgiunte d'altra parte provoca altre complicazioni, come quella di dover unificare più letterali contemporaneamente. Se si considera $\{P(x) \vee P(y), \neg P(z) \vee \neg P(u)\}$, che con una sola risoluzione dà $\Box$, non è possibile refutarlo risolvendo un letterale alla volta; infatti risolvendo ad esempio rispetto a $P(x)$ e $\neg P(z)$ la risolvente non sarebbe $P(y) \vee \neg P(u)$, che risolta con la prima, con $\{x/u\}$, darebbe $P(y)$ che poi sarebbe eliminato con due risoluzioni con la seconda, ma sarebbe $P(x_1) \vee \neg P(x_2)$, ancora con due letterali; solo unificando i due letterali positivi con uno negativo, o con entrambi i negativi, si arriva alla clausola vuota; da qui la formulazione della regola.

Tuttavia nell'esecuzione manuale si può talvolta evitare di applicare alla risolvente la rinomina, quando sia certo che i letterali che vi compaiono non dovranno o potranno più essere risolti con letterali di clausole già presenti a variabili non disgiunte.

Si potrebbe evitare la rinomina supponendo di avere a disposizione tante copie diverse delle clausole iniziali, cosa che è sempre possibile perché $\forall x C$ è equivalente a $\forall x C \wedge \forall y C[x/y]$, copie che si potrebbero produrre al momento del bisogno; non si risolverebbe tuttavia il problema dell'eventuale risoluzione con clausole precedentemente ottenute.

Valgono le proprietà di correttezza e completezza per la risoluzione con variabili: un insieme di clausole è insoddisfacibile se e solo se da esso è derivabile la clausola vuota.

Per la dimostrazione della completezza, un ruolo essenziale gioca il seguente

Lemma 8.3.10 (Lifting)**.** *Se $C_1 \cup C_2$ è la risolvente proposizionale di due clausole chiuse $C_1 \cup \{L\}$ e $C_2 \cup \{L^c\}$ provenienti, con due sostituzioni θ_1 e θ_2, dalle clausole con variabili disgiunte*

$$C_1' \cup \{L(t_{11}, \ldots, t_{1n}), \ldots, L(t_{h1}, \ldots, t_{hn})\}$$

e

$$C_2' \cup \{L^c(s_{11}, \ldots, s_{1n}), \ldots, L^c(s_{k1}, \ldots, s_{kn})\}$$

dove

$$\{L(t_{11}, \ldots, t_{1n}), \ldots, L(t_{h1}, \ldots, t_{hn})\}\theta_1 = \{L\}$$

e

$$L^c(s_{11}, \ldots, s_{1n}), \ldots, L^c(s_{k1}, \ldots, s_{kn})\}\theta_2 = \{L^c\},$$

allora se $C_1'\sigma \cup C_2'\sigma$ è la risolvente di queste due clausole, esiste una sostituzione λ tale che $(C_1'\sigma \cup C_2'\sigma)\lambda = C_1 \cup C_2$.

Dimostrazione. Se D_1' e D_2' sono le due clausole con variabili a cui sono applicate θ_1 e θ_2, l'enunciato del *lifting* è riassunto dal seguente diagramma:

$$\begin{tikzcd}
D_1', D_2' \arrow[r, dashed, "\sigma"] \arrow[d, "\theta_1" ', "\theta_2"] & C_1'\sigma \cup C_2'\sigma \arrow[d, dashed, "\lambda"] \\
C_1 \cup \{L\}, C_2 \cup \{L^c\} \arrow[r] & C_1 \cup C_2
\end{tikzcd}$$

La dimostrazione si basa sulla semplice osservazione che, se σ è un unificatore *generale* di

$$\{L(t_{11},\ldots,t_{1n}),\ldots,L(t_{h1},\ldots,t_{hn}),L(s_{11},\ldots,s_{1n}),\ldots,L(s_{k1},\ldots,s_{kn})\}$$

allora dopo aver considerato che θ, l'unione di θ_1 e θ_2 che agiscono su variabili disgiunte, è anch'essa un unificatore dello stesso insieme, deve esistere λ tale che $\theta = \sigma\circ\lambda$. $\square$

Valgono i raffinamenti analoghi a quelli della risoluzione proposizionale, ad esempio la risoluzione lineare ordinata e, per le clausole di Horn, la risoluzione a input.

In pratica, per trovare una refutazione di un insieme $\mathcal{S}$ di clausole, conviene cercare prima una derivazione proposizionale della clausola vuota dalle clausole ottenute in questo modo: dalle formule atomiche occorrenti nelle clausole si cancellano i termini, lasciando solo il simbolo predicativo che è come una lettera proposizionale. Trovata una derivazione proposizionale si vede se la stessa struttura può essere mantenuta ripristinando le formule originarie e accompagnando le risoluzioni con le necessarie unificazioni.

Esempio 8.3.11.

L'insieme $\mathcal{S} = \{R(c,x) \vee \neg P(y), \neg R(x,F(x)), P(F(x)) \vee R(x,y)\}$ è la matrice dell'enunciato

$$\forall x \exists z \forall y((P(y) \to R(c,x)) \wedge (\neg P(z) \to R(x,y)) \wedge \neg R(x,z)).$$

Si consideri l'insieme associato di clausole proposizionali $\{R\vee\neg P, \neg R, P\vee R\}$; esso ammette la refutazione

$$\begin{array}{cccc}
R \vee \neg P & \neg R & & \\
\searrow & \swarrow & & \\
& \neg P & P \vee R & \\
& \searrow & \swarrow & \\
& & R & \neg R \\
& & \searrow & \swarrow \\
& & & \square
\end{array}$$

che è una guida per la seguente refutazione di S, dopo che si sono riscritte le clausole come $\{R(c,x)\vee\neg P(y), \neg R(z,F(z)), P(F(v))\vee R(v,w)\}$, a variabili disgiunte:

$$\begin{array}{ccc}
R(c,x) \vee \neg P(y) & \neg R(z,F(z)) & \\
\searrow & \swarrow_{\{z/c,x/F(c)\}} & \\
& \neg P(y) & P(F(v)) \vee R(v,w) \\
& \searrow & \swarrow_{\{y/F(v)\}}
\end{array}$$

$$
\begin{tikzcd}[column sep=small]
R(v,w) \arrow[dr] & & \neg R(z,F(z)) \arrow[dl, "\{y/z,\, w/F(z)\}"] \\
& \square &
\end{tikzcd}
$$

Si noti che se esiste una derivazione per risoluzione con variabili, come l'ultima, allora cancellando tutti i termini e le sostituzioni resta la derivazione proposizionale precedente. L'esistenza di questa è dunque condizione necessaria per l'esistenza di quella con variabili.

Non è vero il viceversa, la condizione non è sufficiente; se si considera il nuovo insieme di clausole $\{R(c,x) \lor \neg P(y), \neg R(F(z),F(z)), P(F(v)) \lor R(v,w)\}$, ad esso è associato lo stesso insieme di clausole proposizionali, con le stesse refutazioni. Ma non è possibile eseguire neanche la prima risoluzione, perché $R(c,x)$ e $R(F(z),F(z))$ non sono unificabili.

Anche alle altre possibili refutazioni dell'insieme di clausole proposizionali non è possibile far corrispondere una refutazione delle clausole con variabili, perché prima o poi si arriva a clausole non unificabili. Ad esempio a

$$
\begin{tikzcd}
P \lor R \arrow[dr] & & \neg R \arrow[dl] & & \\
& P \arrow[dr] & & R \lor \neg P \arrow[dl] & \\
& & R \arrow[dr] & & \neg R \arrow[dl] \\
& & & \square &
\end{tikzcd}
$$

corrisponde

$$
\begin{tikzcd}
P(F(v)) \lor R(v,w) \arrow[dr] & & \neg R(F(z),F(z)) \arrow[dl, "\{v/F(z),w/F(z)\}"] & \\
& P(F(F(z))) \arrow[dr] & & R \lor \neg P \arrow[dl, "\{y/F(F(z))\}"] \\
& & R(c,x) &
\end{tikzcd}
$$

che si interrompe qui perché $R(c,x)$ non è unificabile con $R(F(z),F(z))$.

Alla stessa conclusione si arriva esplorando le altre possibilità (esercizio). In effetti il nuovo insieme di clausole è la matrice dell'enunciato soddisfacibile $\forall x \exists z \forall y((P(y) \to R(c,x)) \land (\neg P(z) \to R(x,y)) \land \neg R(z,z))$.

Esercizi

Esercizio 8.3.12. Verificare con la risoluzione se i seguenti insiemi di clausole sono insoddisfacibili:

$$\{P(x), \neg P(x) \lor Q(F(x)), \neg Q(F(c))\}$$
$$\{P(x), Q(x, F(x)) \lor \neg P(x), \neg Q(G(y), z)\}$$
$$\{P(x) \lor Q(F(x)), \neg P(c) \lor R(x, y), \neg R(c, x), \neg Q(F(c))\}$$
$$\{\neg P(x), \neg R(F(x), y), P(F(x)) \lor R(x, c) \lor Q(x), \neg Q(x) \lor P(y)\}$$
$$\{\neg P(x) \lor Q(x), \neg Q(c) \lor \neg P(x), P(x) \lor R(F(x)), \neg R(x)\}.$$

Esercizio 8.3.13. Dimostrare con la risoluzione che

$$\{\neg P(x), \neg R(F(x), y), P(F(x)) \lor R(x, c) \lor Q(x), \neg Q(x) \lor P(y)\}$$

è insoddisfacibile.

Esercizio 8.3.14. Dimostrare che

$$\{\neg P(x) \lor R(x, F(x)), \neg R(c, x) \lor Q(x),$$
$$\neg Q(G(x)) \lor P(x), P(c) \lor Q(x), \neg P(x) \lor \neg Q(F(x))\}$$

è insoddisfacibile, mentre

$$\{\neg P(x) \lor R(x, F(x)), \neg R(c, x) \lor Q(x), \neg Q(G(x)) \lor P(G(x)),$$
$$P(c) \lor Q(x), \neg P(x) \lor \neg Q(F(x))\}$$

non lo è.

Esercizio 8.3.15. Verificare, riconducendosi al calcolo della risoluzione, se

$$\forall x \exists y (P(x, y) \lor Q(x, y)), \forall x \forall y (P(x, y) \to R(x)),$$
$$\forall x \forall y (Q(x, y) \to R(x)) \models \exists x R(x)$$

$$\forall x P(x), \exists x P(x) \to \forall x \exists y R(x, y) \models \forall x \exists y R(x, y)$$

$$\exists x \neg P(x) \to \exists y \neg R(y), \forall x P(x) \to \exists x Q(x),$$
$$\exists x Q(x) \to \neg \forall x R(x) \models \exists x \neg R(x)$$

$$\exists x P(x) \to \forall x R(c, x), \forall x (\forall z \neg P(z) \to \forall v R(x, v)) \models \exists x \forall y R(x, y).$$

8.4 Programmazione logica

Esempio 8.4.1. Sia data la seguente storia: Giovanni ama Maria? Se uno si emoziona quando guarda una ragazza, allora si può dire che l'ama. Giovanni si emoziona se sta vicino a Maria e Maria sorride. Uno sorride quando è contento. Se a uno fanno un regalo che gli piace, questi è contento. A Maria piaccioni i dischi di Baglioni. Giovanni regala a Maria un disco di Baglioni e sta sempre vicino a Maria e non le toglie gli occhi di dosso.

Con una opportuna semplificazione e formalizzazione, la storia può essere resa dalla congiunzione dei seguenti enunciati:

$$\begin{aligned}
&\forall x \forall y (G(x,y) \land E(x) \to (R(y) \to A(x,y)))\\
&R(m)\\
&V(g,m) \land S(m) \to E(g)\\
&\forall x (C(x) \to S(x))\\
&\forall x \forall z (\exists y (R(x,y,z) \land P(y,z)) \to C(z))\\
&\forall x (D(x) \to P(x,m))\\
&\exists x (D(x) \land R(g,x,m))\\
&V(g,m)\\
&V(g,m) \to G(g,m),
\end{aligned}$$

ovvero in forma di clausole:

$$\begin{array}{ll}
1 & A(x,y) \leftarrow R(y), G(x,y), E(y)\\
2 & R(m) \leftarrow\\
3 & E(g) \leftarrow V(g,m), S(m)\\
4 & S(z) \leftarrow C(z)\\
5 & C(w) \leftarrow R(u,v,w), P(v,w)\\
6 & P(j,m) \leftarrow D(j)\\
7 & D(d) \leftarrow\\
8 & R(g,d,m) \leftarrow\\
9 & V(g,m) \leftarrow\\
10 & G(g,m) \leftarrow V(g,m),
\end{array}$$

con il *goal*

$$\begin{array}{ll}
11 & \leftarrow A(g,m),
\end{array}$$

a cui si risponde con la seguente derivazione:

12	$\leftarrow R(m), G(g,m), E(g)$	da 1 con $\{x/g, y/m\}$
13	$\leftarrow G(g,m), E(g)$	da 2
14	$\leftarrow V(g,m), E(g)$	da 10
15	$\leftarrow E(g)$	$da9$
16	$\leftarrow V(g,m), S(m)$	da 3
17	$\leftarrow S(m)$	da 9
18	$\leftarrow C(m)$	da 4 con $\{z/m\}$
19	$\leftarrow R(u,v,m), P(v,m)$	da 5 con $\{w/m\}$
20	$\leftarrow P(d,g)$	da 8 con $\{u/g, v/d\}$
21	$\leftarrow D(d)$	da 6 con $\{j/d\}$
22	$\leftarrow$	da 7

e il *goal* è soddisfatto.

Esempio 8.4.2. Sia dato il programma

1 $\quad P(x) \leftarrow$

2 $\quad R(c,y) \leftarrow P(y)$

con il *goal*

3 $\quad \leftarrow R(z, F(z)),$

Ad esso si risponde unificando il *goal* con la testa della 2, ottenendo come nuovo *goal*

4 $\quad \leftarrow P(F(c))$

e quindi

5 $\quad \leftarrow$

con la sostituzione $\{x/F(c)\}$.

Si noti che in una reale applicazione della programmazione logica la risposta che sarebbe data è

$$z := c$$

in quanto la richiesta non è tanto se il *goal* è soddisfatto, ma da chi. In notazione logica, l'interrogazione dell'esempio deriva dal desiderio di sapere se

$$\forall x P(x), \forall x(P(x) \to R(c,x)) \models \exists x R(x, F(x)).$$

Le variabili del *goal* derivano da quantificatori esistenziali, e richiedono per la risposta positiva un'esemplificazione specifica; se il *goal* deriva da una putativa conseguenza universale del programma, allora non ha variabili ma costanti, come effetto della trasformazione mediante la negazione.