

# Logica

# PRESENTAZIONE

## SCOPO:

Formalizzare i meccanismi di ragionamento

## ORIGINI:

Più di duemila anni fa (filosofi greci e indiani)

## LOGICA MODERNA:

Nata tra la fine dell'Ottocento e l'inizio del  
Novecento (Boole, Peano, Frege, Russel, ...)

# LA LOGICA STUDIA *PROPOSIZIONI*, E CIOÈ ESPRESSIONI CHE RAPPRESENTANO *AFFERMAZIONI*

Nella formulazione standard una proposizione può essere VERA o FALSA, ma non ambedue (si tratta cioè di una logica a due valori).

È necessario stabilire:

- che cosa si vuole formalizzare
- in che modo scrivere le proposizioni (**Sintassi**)
- come determinare il significato di una proposizione (**Semantica**)

# Logica Proposizionale

$P_1$ : Se fa caldo ed è umido, allora pioverà

$P_2$ : Se è umido ed è estate, allora fa caldo

$P_3$ : Adesso è umido

$P_4$ : Adesso è estate

Si vuole verificare:

$P_5$ : Pioverà

In logica proposizionale, ad ogni proposizione elementare viene associata un *variabile proposizionale*

A = FA CALDO

C = È ESTATE

B = È UMIDO

D = PIOVERÀ

La rappresentazione dell'esempio è:

$$F_1: A \wedge B \rightarrow D$$

$$F_2: B \wedge C \rightarrow A$$

$$F_3: B$$

$$F_4: C$$

in cui  $\wedge$  è il simbolo che rappresenta la *coniunzione* (AND) e  $\rightarrow$  è il simbolo che rappresenta la *implicazione*.

Si vuole dimostrare che da  $F_1$ - $F_4$  *segue logicamente*  $F_5: D$

Intuitivamente, si ha:

Step 1: Se adesso è umido ( $P_3$ ) ed è estate ( $P_4$ ), allora si può dedurre che fa caldo

Step 2: Se adesso fa caldo (da Step 1) ed è umido ( $P_3$ ), allora da  $P_1$  si può dedurre che pioverà.



Il problema della logica proposizionale è che le proposizioni elementari sono unità atomiche non scomponibili.

Si consideri:

$P_6$ : Il 18/8/90 qui faceva caldo

$P_7$ : Qui fa sempre caldo

$P_8$ : C'è almeno un giorno dell'anno in cui qui fa caldo

Queste proposizioni non sono indipendenti (ad esempio, se  $P_7$  è vera,  $P_6$  non può essere falsa), ma non vi è modo in logica proposizionale di esprimere questa dipendenza reciproca.

La dipendenza è dovuta alla presenza dello stesso **predicato** ("*fare caldo*") nelle tre proposizioni.

Per esaminare la **struttura interna** delle proposizioni è necessario uno strumento più potente: la **logica dei predicati**.

# Logica dei predicati

Introduzione di **individui e variabili individuali**, di **funzioni e predicati**

$P_6$  (prima indicato semplicemente da una variabile proposizionale - es. E) può essere scritto:

$P_6: \text{CALDO}(18/8/90)$

in cui CALDO è un *predicato* e 18/8/90 è un *individuo*

Vengono introdotti i **quantificatori**

$\exists$ : esiste

$\forall$ : per ogni

$P_7: (\forall x)CALDO(x)$

$P_8: (\exists x)CALDO(x)$

$x$  è una **variabile individuale**

Ora da  $P_7$  si può dedurre  $P_6$ , perché 18/8/90 è uno dei possibili "valori" (detto impropriamente) di  $x$ , e il predicato che appare nelle due formule è lo stesso.

La logica dei predicati è **indecidibile**.

Ciò significa che non esiste nessun algoritmo che permetta *sempre* di determinare se una data proposizione segue logicamente da un insieme di proposizioni dato.

È però possibile progettare algoritmi tali che, se la proposizione segue logicamente, riescono a determinarlo, ma se ciò non è vero, possono operare all'infinito.

Si noti, comunque, che tali risultati valgono per domini infiniti.

Scopo di molte ricerche attuali è di progettare algoritmi che operino in modo *efficiente*.

Nel corso si parlerà di due metodi: la **deduzione naturale** e il **metodo di risoluzione**.

Alcuni vantaggi possono essere ottenuti ponendo dei vincoli sulla struttura delle proposizioni. Questo è il caso delle **Clausole di Horn**, che sono alla base del linguaggio PROLOG.

In alcuni casi, è necessario introdurre *estensioni* della logica dei predicati.



Uno dei principi fondamentali della logica dei predicati è che, se in una formula vera si sostituisce ad una sua parte qualcosa di equivalente, si ottiene ancora una formula vera.

Questo non è sempre corretto.

$P_9$ : Giorgio conosce il numero di telefono di Maria

$P_{10}$ : Il numero di telefono di Maria è uguale al numero di telefono di Carlo

Da  $P_9$  e  $P_{10}$  è errato dedurre:

$P_{11}$ : Giorgio conosce il numero di telefono di Carlo

Questo perché "conoscere" è un operatore diverso dagli altri: è un operatore **modale**.

# La logica proposizionale

La logica proposizionale tratta *formule*

Una formula è composta da

- *formule atomiche* (o atomi)
- *connettivi logici*

Esempi di formule atomiche sono A, B, C, D viste in precedenza.

I connettivi logici più comunemente usati sono:

$\neg$  (NOT: negazione)

$\vee$  (OR: disgiunzione)

$\wedge$  (AND: congiunzione)

$\rightarrow$  (IMPLIES o IF...THEN...: implicazione)

$\leftrightarrow$  (IF AND ONLY IF: bi-implicazione)

Una formula è *ben formata* (FBF) se e solo se  
essa è ottenibile applicando le seguenti regole:

1. un atomo è una FBF

2. se  $F$  è una FBF, allora  $(\neg F)$  è una FBF

3. se  $F$  e  $G$  sono FBF, allora  $(F \vee G)$ ,  $(F \wedge G)$ ,  
 $(F \rightarrow G)$ ,  $(F \leftrightarrow G)$  sono FBF

Esempi:

FBF:

$$((P \vee Q) \rightarrow R)$$

$$((P \vee Q) \leftrightarrow (\neg((\neg P) \wedge (\neg Q))))$$

Non FBF:

$$(P \neg(Q \rightarrow))$$

$$((P \wedge Q \wedge R) \rightarrow Q \vee S)$$

Le regole viste esprimono la SINTASSI (vincoli strutturali) delle formule del calcolo proposizionale.

Stabilendo un ordinamento tra i connettivi è possibile eliminare alcune parentesi.

L'ordine che verrà adottato è il seguente:

1.  $\neg$

2.  $\wedge, \vee$

3.  $\rightarrow, \leftrightarrow$

La formula

$$((((A \vee B) \wedge C) \rightarrow D) \leftrightarrow (((\neg D) \rightarrow A) \vee E))$$

può essere riscritta (eliminando le parentesi esterne):

$$((A \vee B) \wedge C \rightarrow D) \leftrightarrow (\neg D \rightarrow A) \vee E$$

Si noti che

$$(A \rightarrow B) \vee C$$

$$A \rightarrow B \vee C$$

Sono due formule distinte.



La SEMANTICA della logica proposizionale richiede l'introduzione dei *valori di verità*

L'insieme dei valori di verità (che indicheremo con  $B$ ) include VERO e FALSO, rappresentati da T (true) e F (false):

$$B = \{T, F\}$$

Una interpretazione  $I$  consiste in un mapping tra l'insieme delle formule e  $B$  (specificando cioè, per ogni formula, se essa è vera o falsa).

Problema:

l'insieme delle formule è INFINITO

Come si può specificare una interpretazione?

Osservazione:

un *assegnamento di valore* alle formule atomiche  
identifica univocamente un'interpretazione.

Soluzione:

specificare i valori di verità delle formule atomiche. Da essi è possibile ricavare il valore di verità di ogni FBF.

Infatti, valgono le seguenti regole, associate ai diversi connettivi:

1. Se  $P$  è vera, allora  $\neg P$  è falsa, e viceversa. Questo si può rappresentare con la tabella:

$P$	$\neg P$
T	F
F	T

2. Se  $P$  è vera e  $Q$  è vera, allora  $(P \wedge Q)$  è vera; in tutti gli altri casi è falsa. In tabella:

$P$	$Q$	$(P \wedge Q)$
T	T	T
T	F	F
F	T	F
F	F	F

3. Se  $P$  è falsa e  $Q$  è falsa, allora  $(P \vee Q)$  è falsa; in tutti gli altri casi è vera. In tabella:

$P$	$Q$	$(P \vee Q)$
T	T	T
T	F	T
F	T	T
F	F	F

4. Se  $P$  è vera e  $Q$  è falsa, allora  $(P \rightarrow Q)$  è falsa; in tutti gli altri casi è vera(\*).

P	Q	$(P \rightarrow Q)$
T	T	T
T	F	F
F	T	T
F	F	T

(\*) La prima affermazione (che corrisponde alla seconda riga della t. di v.) è evidente: corrisponde al significato dell'affermazione  
"If P then Q".

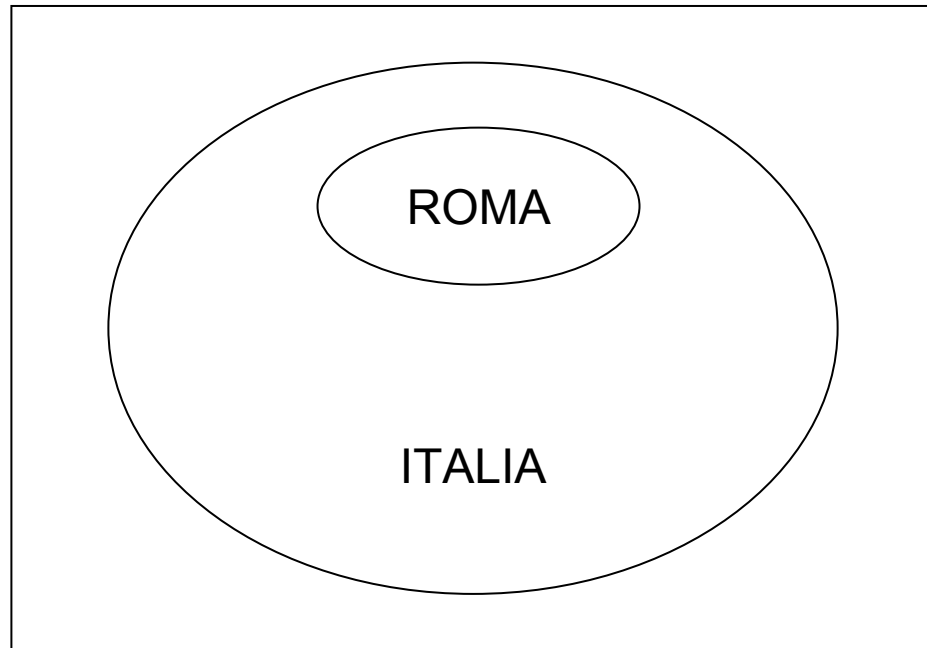
Un modo per decidere come riempire il resto della tavola di verità è quello di considerare la proposizione "If  $(C \wedge D)$  then C", che evidentemente è sempre vera.

Quando C è T e D è F,  $(C \wedge D)$  è F.

Allora la terza riga della t. di v. deve essere T (poiché  $(C \wedge D)$  è F, C è T, e "If  $(C \wedge D)$  then C" è T.

Analogamente si ragiona per le altre due righe.

In alternativa, si può ragionare con i diagrammi di Venn: si consideri l'affermazione "If sono\_a\_Roma then sono\_in\_Italia". Se sono\_a\_Roma è T e sono\_in\_Italia è F, l'affermazione è F, in tutti gli altri casi è T.





5. Se P e Q hanno lo stesso valore di verità, allora  $(P \leftrightarrow Q)$  è vera, altrimenti è falsa. In tabella:

P	Q	$(P \leftrightarrow Q)$
T	T	T
T	F	F
F	T	F
F	F	T

Vediamo come è possibile determinare il valore di verità di una formula in base agli *assegnamenti di valore* delle formule atomiche.

Esempio:  $(P \wedge Q) \vee R \rightarrow (P \leftrightarrow R) \wedge Q$

Ponendo:

$$\alpha = (P \wedge Q) \vee R$$

$$\beta = (P \leftrightarrow R) \wedge Q$$

P	Q	R	$(P \wedge Q)$	$\alpha$	$(P \leftrightarrow R)$	$\beta$	$\alpha \rightarrow \beta$
T	T	T	T	T	T	T	T
T	T	F	T	T	F	F	F
T	F	T	F	T	T	F	F
T	F	F	F	F	F	F	T
F	T	T	F	T	F	F	F
F	T	F	F	F	T	T	T
F	F	T	F	T	F	F	F
F	F	F	F	F	T	F	T

Una tabella come questa è detta **tavola di verità**.

Intuitivamente, ogni riga di una tabella di verità corrisponde ad una diversa possibile situazione (interpretazione)

Alcune formule sono vere in tutte le interpretazioni.

Esempio:  $(P \wedge (P \rightarrow Q)) \rightarrow Q$

P	Q	$P \rightarrow Q$	$P \wedge (P \rightarrow Q)$	$(P \wedge (P \rightarrow Q)) \rightarrow Q$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

Esse sono dette **formule valide** o **tautologie**.

Altre formule sono false in tutte le interpretazioni.

Esempio:  $(P \rightarrow Q) \wedge P \wedge \neg Q$

P	Q	$P \rightarrow Q$	$\neg Q$	$(P \rightarrow Q) \wedge P \wedge \neg Q$
T	T	T	F	F
T	F	F	T	F
F	T	T	F	F
F	F	T	T	F

Esse sono dette **formule inconsistenti** o **contraddizioni**.

Poiché ogni formula è finita e quindi contiene un numero finito di formule atomiche, è sempre possibile determinare se essa è valida, inconsistente o né l'una né l'altra.

La logica proposizionale è quindi **decidibile**.

# I paradossi dell'implicazione materiale

Alcune formule valide sono controintuitive. In esse compare il connettivo  $\rightarrow$

Esempi: 1)  $\neg P \rightarrow (P \rightarrow Q)$

P	Q	$\neg P$	$P \rightarrow Q$	$\neg P \rightarrow (P \rightarrow Q)$
T	T	F	T	T
T	F	F	F	T
F	T	T	T	T
F	F	T	T	T

$$2) Q \rightarrow (P \rightarrow Q)$$

$$3) (P \rightarrow Q) \vee (P \rightarrow \neg Q)$$

$$4) (P \rightarrow Q) \vee (Q \rightarrow P)$$

Si ribadisce che l'implicazione non ha nulla a che vedere con la *causalità*.



Due formule F e G sono **equivalenti** (scritto  $F \Leftrightarrow G$ ) se e solo se esse hanno lo stesso valore di verità in tutte le interpretazioni.

Esempi:

$$1) P \rightarrow Q \Leftrightarrow \neg P \vee Q$$

P	Q	$P \rightarrow Q$	$\neg P$	$\neg P \vee Q$
T	T	T	F	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

$$2) \quad a - P \vee Q \Leftrightarrow Q \vee P$$

$$b - P \wedge Q \Leftrightarrow Q \wedge P$$

Leggi commutative

$$3) \quad a - (P \vee Q) \vee R \Leftrightarrow P \vee (Q \vee R)$$

$$b - (P \wedge Q) \wedge R \Leftrightarrow P \wedge (Q \wedge R)$$

Leggi associative

$$4) \quad a - P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R)$$

$$b - P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$$

Leggi distributive

$$5) \quad a - \neg(P \vee Q) \Leftrightarrow \neg P \wedge \neg Q$$

$$b - \neg(P \wedge Q) \Leftrightarrow \neg P \vee \neg Q$$

Leggi di De Morgan

Le leggi associative permettono di eliminare le parentesi in caso di sequenze di  $\wedge$  o di  $\vee$

Esempio:

$$((((F \vee G) \vee H) \vee L) \vee (M \vee N))$$

può essere riscritta come

$$F \vee G \vee H \vee L \vee M \vee N$$

Si osservi che "  $F \vee G \vee H$  " non è una FBF, ma solo una abbreviazione, per di più ambigua. Essa infatti rappresenta sia

$$((F \vee G) \vee H)$$

che

$$(F \vee (G \vee H))$$

Questa ambiguità non causa problemi perché le due formule sono equivalenti.

La proprietà associativa **non vale** per l'implicazione:

$$(F \rightarrow (G \rightarrow H)) \not\leftrightarrow ((F \rightarrow G) \rightarrow H)$$

(dove ! sta per NOT)

Non sarà quindi permesso scrivere

$$F \rightarrow G \rightarrow H$$

# La logica: formalizzazione

Si possono rappresentare fatti del mondo mediante proposizioni logiche scritte come formule ben formate (fbf).

Esempi

PIOVE si esprime un fatto (piove)

SOLE si esprime un fatto (c'è il sole)

$\text{PIOVE} \rightarrow \neg \text{SOLE}$  si esprime il fatto che se piove, non c'è il sole.

La logica proporzionale è limitata.

Esempio:

Socrate è un uomo:

SOCRATE UOMO

Anche Platone è un uomo

PLATONE UOMO

I due fatti non trovano alcuna correlazione.

Nella logica dei predicativi si supera questo handicap:

uomo (Socrate)

uomo (Platone)

in generale:

uomo (x)



Per l'uso del calcolo dei predicati occorre definire la sintassi e la semantica (perché si tratta di un “linguaggio” per modellare il mondo).

Sintassi:

- specifica un alfabeto di simboli
- definisce le espressioni che possono essere costruite

L'alfabeto è costituito da:

1)simboli di interpunzione: , ( )

2)simboli logici o connettivi logici:  $\neg$ ,  $\vee$ ,  $\wedge$ ,  $\rightarrow$ ,  
 $\leftrightarrow$

3. lettere funzionali n-adiche  $f_i^n$  ( $i \geq 1, n \geq 0$ ).  
Si usa  $f_i^n$  per avere un set arbitrariamente grande di lettere funzionali.  
Si usa a, b, c al posto di  $f_i^0$  (lettere costanti),  
f, g, h per  $f_i^1$  (funzioni).

4. lettere predicative n-adiche:  
 $p_i^n$  ( $i \geq 1, n \geq 0$ )

Nella pratica, si usano P, Q, R (predicati)

Mediante questi simboli vengono costruite varie espressioni, che si possono definire ricorsivamente come:

### 1. Termini:

1. ogni lettera costante è un termine;
2. se  $t_1, t_2, \dots, t_n$  ( $n \geq 1$ ) sono termini, anche  $f_i^n(t_1, t_2, \dots, t_n)$  lo è;
3. nessun'altra espressione è un termine.

Si noti che quando  $g(t_1, t_2, \dots, t_n)$  è usata come termine, essa sta al posto di una  $f_i^n(t_1, t_2, \dots, t_n)$ ;  $g$  non ha bisogno di superscritte.

## 2. Formule atomiche:

1. le lettere preposizionali sono formule atomiche;
2. se  $t_1, t_2, \dots, t_n$  ( $n \geq 1$ ) sono termini, l'espressione  $p_i^n(t_1, t_2, \dots, t_n)$  è una forma atomica;
3. nessun'altra espressione è un formula atomica.

### 3. Formule ben formate (fbff):

1. ogni formula atomica è una fbf;
2. se  $A$  è una fbf, anche  $(\neg A)$  lo è;
3. se  $A$  e  $B$  sono fbff, anche  $(A \rightarrow B)$  lo è;
4. nessun'altra espressione è una fbf.

(Si osservi che alcuni studiosi fanno discendere l'AND e l'OR dall'implicazione: quindi anche  $(A \wedge B)$  e  $(A \vee B)$  sono fbff).

Esempi di fbff:

$$\neg P(a, g(a, b, a))$$

$$P(a, b) \rightarrow (\neg Q)$$

$$(\neg(P(a) \rightarrow P(b))) \rightarrow P(b)$$

$$\neg(P(a) \rightarrow Q(f(a)))$$

e degli esempi di espressioni che non sono fbff:

$$\neg f(a)$$

$$f(P(a))$$

$$Q(f(a), (P(b) \rightarrow P(c)))$$

(In buona sostanza, sono fbff quelle che sono riconducibili a Vero o Falso).

Semantica:

una fbf assume un “significato” quando è interpretata come affermazione sul dominio del discorso.

Dominio: insieme non vuoto (anche infinito).

Es.: insieme degli interi, insieme degli uomini, ecc.



Semantica:

una fbf assume un “significato” quando è interpretata come affermazione sul dominio del discorso.

Dominio: insieme non vuoto (anche infinito).

Es.: insieme degli interi, insieme degli uomini, ecc.

Le affermazioni evidenziano relazioni fra gli elementi del dominio:

Padre (Mario, Piero)

Funzioni sul domino:

Dato un dominio  $D$ , una funzione  $n$ -adica manda ogni  $n$ -pla di elementi di  $D$  in un elemento di  $D$ .

Es.: La funzione Più manda coppie di interi in altri interi (secondo le regole dell'addizione).

Si definisce un “interpretazione” ovvero un “modello” di una fbf se, oltre a fissare il dominio  $D$ ,

1. ad ogni simbolo costante della fbf associamo un particolare elemento di  $D$ ;
2. ad ogni lettera funzionale della fbf associamo una particolare funzione di  $D$  (a lettere funzionale  $n$ -adiche corrispondono funzioni  $n$ -adiche);
3. ad ogni lettera predicativa della fbf associamo una particolare relazione fra gli elementi di  $D$  (a lettere predicative  $n$ -adiche corrispondono relazioni  $n$ -adiche).

Data una fbf e una interpretazione, ogni formula atomica assume un valore Vero o Falso (se i termini della lettera predicativa corrispondono ad elementi di D che soddisfano la relazione associata, il valore è V).

Esempio:

$$P(a, f(b, c))$$

e la seguente interpretazione:

$D$  è l'insieme degli interi

$a$  è l'intero 2

$b$  è l'intero 4

$c$  è l'intero 6

$f$  è la funzione d'addizione

$P$  è la relazione *maggiore di*

Il significato della formula atomica è

$2$  è *maggiore di* ( $4$  più  $6$ )

e il valore è Falso.

Variabili individuali  $x_i$ : sono usate per indicare una generica costante.

Quantificatore universale: esprime il concetto di “per tutti i valori assunti”. Si indica con  $\forall$

Esempio:

$$(\forall x) P(x)$$

$x$  è detta “variabile universalmente quantificata”

Quantificatore esistenziale: esprime il concetto di “esiste almeno un elemento di D”. Si indica con  $\exists$ .

Esempio:

$$(\exists x) P(x)$$

x è detta “esistenzialmente quantificata”.

Si può dimostrare, per domini finiti, che:

$$\neg(\forall x) W(x)$$

è equivalente a

$$(\exists x) \{\neg W(x)\}$$

e che:

$$\neg(\exists x) W(x)$$

è equivalente a

$$(\forall x) \{\neg W(x)\}$$

(basta costruire la tavola di verità)

Si estendono queste proprietà a domini infiniti.

N.B.: è l'estensione di De Morgan



# Il ragionamento

Uso della logica dei predicati: determinare la validità di una proposizione, date certe premesse.

Nota: la logica proposizionale è decidibile (basta applicare la tavola di verità).

La logica dei predicati NON è decidibile.

Esistono procedure di prova di un teorema, se questo è un teorema, altrimenti quelle procedure possono non terminare.

Nonostante ciò, trovano utili applicazioni.

Vediamo un metodo empirico.

Siano dati questi fatti:

1. uomo (Marco)

Marco era un uomo (la temporalità qui è omessa)

2. abitante\_di\_Pompei (Marco)

3.  $(\forall x)\text{abitante\_di\_Pompei}(x) \rightarrow \text{romano}(x)$   
tutti gli abitanti di Pompei erano romani

4. Capo (Cesare)

Cesare era un capo

5.  $(\forall x)\text{romano}(x) \rightarrow \text{fedele}(x, \text{Cesare}) \vee \text{odia}(x, \text{Cesare})$   
tutti i romani erano fedeli a Cesare o lo odiavano  
(OR inclusivo, per semplicità)
6.  $\forall x \exists y \text{ fedele}(x, y)$   
Tutti sono fedeli a qualcuno
7.  $\forall x \forall y \text{ persona}(x) \wedge \text{capo}(y) \wedge \text{cerca\_di\_assassinare}(x, y) \rightarrow \neg \text{fedele}(x, y)$   
gli uomini cercano solo di assassinare i capi a cui non sono fedeli.

8. cerca\_di\_assassinare (Marco, Cesare)  
Marco cercò di assassinare Cesare.

Vogliamo rispondere alla domanda:  
Marco era fedele a Cesare?

Ovvero ci prefissiamo di dimostrare la verità di  
 $\neg$ Fedele (Marco, Cesare)

Si può partire dagli assunti, effettuare le sostituzioni e vedere se si arriva alla meta: alto fattore di ramificazione.

In alternativa: si ragiona all'indietro, partendo dalla meta (basso fattore di ramificazione)

Per provare la meta, si usano le regole di inferenza per trasformarle in altra meta (o altre mete), e così via finché sono soddisfatte tutte le mete (si ottiene un grafo AND-OR).

Esempio (è mostrato un solo ramo)

$\neg \text{Fedele}(\text{Marco}, \text{Cesare})$   
↓ (7, sostituzione)

$\text{persona}(\text{Marco}) \wedge \text{capo}(\text{Cesare}) \wedge$   
 $\text{cerca\_di\_assassinare}(\text{Marco}, \text{Cesare})$   
↓ (4)

$\text{persona}(\text{Marco}) \wedge \text{cerca\_di\_assassinare}(\text{Marco}, \text{Cesare})$   
↓ (8)

$\text{persona}(\text{Marco})$

Non c'è la dimostrazione perché i fatti a disposizione non dicono nulla circa persona (Marco): la meta non può essere soddisfatta.

Se si aggiunge il fatto:

9.  $\forall x \text{ uomo}(x) \rightarrow \text{persona}(x)$

Si arriva al fatto *vero*  $\text{uomo}(\text{Marco})$  e quindi alla dimostrazione.

## Conclusioni:

- occorre disambiguare le frasi espresse in linguaggio naturale
- difficoltà anche nella scelta fra modi diversi di rappresentare la conoscenza
- se la conoscenza non è completa, non si raggiunge la dimostrazione.



Come riconoscere questo fatto?

- quale enunciato è meglio dedurre?

Esempio:

$\text{fedele}(\text{Marco}, \text{Cesare})$

oppure

$\neg \text{fedele}(\text{Marco}, \text{Cesare})?$

(Alcuni sistemi li provano entrambi)!

# Cosa serve per la risoluzione?

Lo desumiamo da alcuni esempi.

Per esprimere fatti

$1 > 0$  ,  $2 > 1$  ecc.

$0 < 1$  ,  $1 < 2$  ecc.

è conveniente disporre di  
*predicati computabili:*

esempio  $>(x, y)$

e di

*funzioni computabili:*

esempio  $> (2+3, 1)$

per il calcolo della somma.

# Vediamo come esprimere alcuni fatti:

1. Marco era un uomo

uomo(Marco)

(si trascura il tempo passato)

2. Marco era un abitante di Pompei:

abitante\_di\_Pompei(Marco)

3. Marco era nato nel 40 d.C.

nato(Marco, 40)

(si trascura d.C., oppure si usano i numeri relativi)

4. Tutti gli uomini sono mortali

$$\forall x \text{ uomo}(x) \rightarrow \text{mortale}(x)$$

5. Tutti gli abitanti di Pompei morirono quando ci fu l'eruzione del vulcano nel 79 d.C.

$$\text{eruzione}(\text{vulcano}, 79) \wedge \forall x (\text{abitante\_di\_Pompei}(x) \rightarrow \text{muore}(x, 79))$$

(nota: concomitanza di fatti e non "causalità")

6. Nessun mortale vive più di 150 anni

$$\forall x \forall t_1 \forall t_2 \text{ mortale}(x) \wedge \text{nato}(x, t_1) \wedge \\ >(t_2 - t_1, 150) \rightarrow \text{morto}(x, t_2)$$

7. Ora siamo nel 1986

$$\text{Ora} = 1986$$

Domanda: Marco è vivo?

Intuitivamente: No perché

- è stato ucciso dal vulcano
- oppure
- avrebbe più di 150 anni

Entrambe le risposte non si possono desumere rigorosamente dalla conoscenza attuale.

Occorre aggiungere:

8. Vivo significa non morto

$$\forall x \forall t \text{ vivo}(x, t) \leftrightarrow \neg \text{morto}(x, t)$$

9. Se qualcuno muore, allora è morto in tutti i momenti successivi

$$\forall x \forall t_1 \forall t_2 \text{ muore}(x, t_1) \wedge >(t_2, t_1) \rightarrow \text{morto}(x, t_2)$$



Riassumendo:

1. uomo(Marco)
2. abitante\_di\_Pompei(Marco)
3. nato(Marco, 40)
4.  $\forall x \text{ uomo}(x) \rightarrow \text{mortale}(x)$
5.  $\forall x (\text{abitante\_di\_Pompei}(x) \rightarrow \text{muore}(x, 79))$
6. eruzione (vulcano, 79)
7.  $\forall x \forall t_1 \forall t_2 \text{ mortale}(x) \wedge \text{nato}(x, t_1) \wedge >(t_2 - t_1, 150) \rightarrow \text{morto}(x, t_2)$
8. Ora = 1986
9.  $\forall x \forall t \text{ vivo}(x, t) \leftrightarrow \neg \text{morto}(x, t)$
10.  $\forall x \forall t_1 \forall t_2 \text{ muore}(x, t_1) \wedge >(t_2, t_1) \rightarrow \text{morto}(x, t_2)$

Si risponde alla domanda:

Marco è vivo?

Dimostrando:

$\neg$ Vivo (Marco, ora)

I modo:

$\neg \text{Vivo}(\text{Marco}, \text{ora})$

(9, sostituzione)

$\downarrow$   
 $\text{morto}(\text{Marco}, \text{ora})$

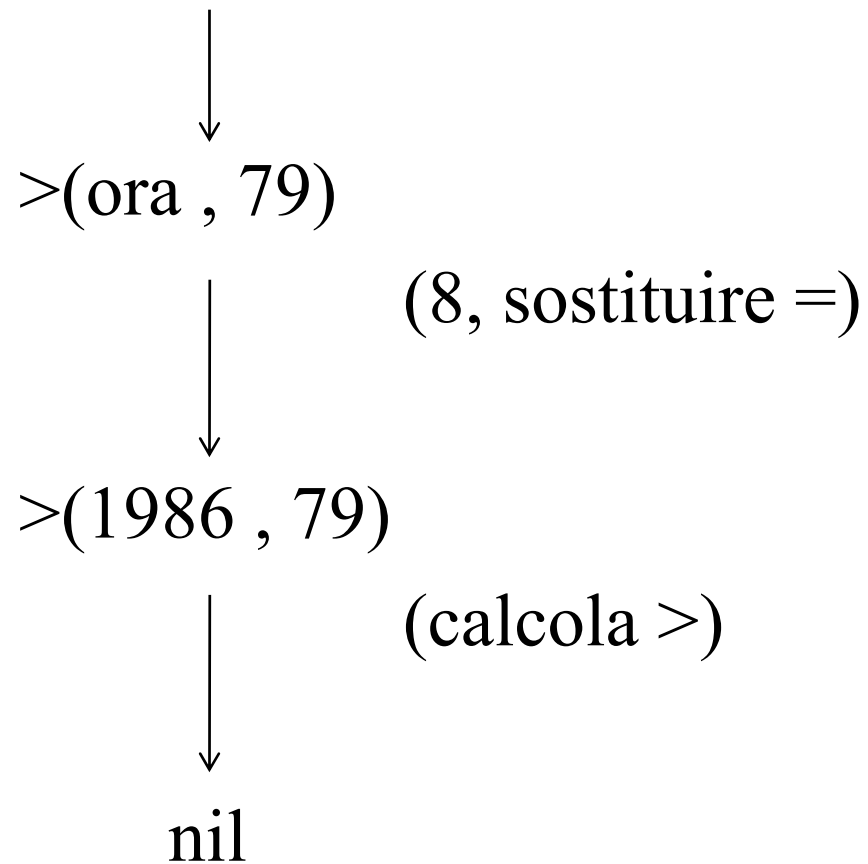
(10, sostituzione)

$\downarrow$   
 $\text{muore}(\text{Marco}, t_1) \wedge >(\text{ora}, t_1)$

(5, sostituzione)

$\downarrow$   
 $(\text{abitante\_di\_Pompei}(\text{Marco}) \wedge >(\text{ora}, 79))$

(2)  $\vdots$



II modo:

$\neg \text{Vivo}(\text{Marco}, \text{ora})$

(9, sostituzione)



$\text{morto}(\text{Marco}, \text{ora})$

(7, sostituzione)



$\text{mortale}(\text{Marco}) \wedge \text{nato}(\text{Marco}, t_1) \wedge >(\text{ora}-t_1, 150)$

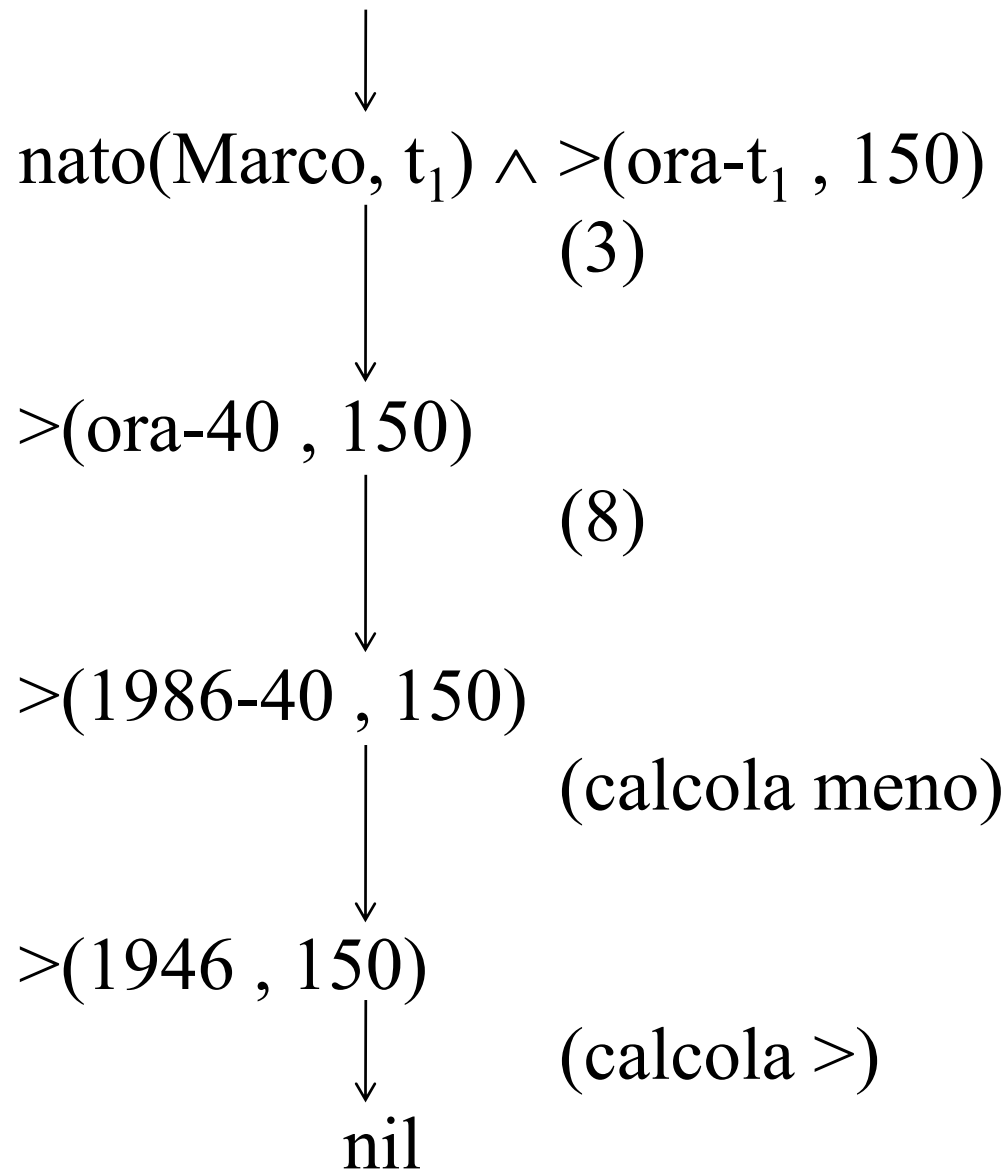
(4, sostituzione)



$\text{uomo}(\text{Marco}) \wedge \text{nato}(\text{Marco}, t_1) \wedge >(\text{ora}-t_1, 150)$



(1)



Nota: il termine NIL significa successo, perché la lista delle condizioni ancora da provare è vuota.

Conclusioni: per arrivare alla dimostrazione occorrono processi quali

L'unificazione (cioè trovare il matching tra le componenti le diverse affermazioni)  
la sostituzione  
l'applicazione del *modus ponens*

Osservazione: le sostituzioni devono essere tutte coerenti. Ad esempio, nella 7), se si pone  $\text{nato}(\text{Marco}, t_1)$  con  $t_1 = 40$ , deve essere anche  $>(\text{ora} - t_1, 150)$  con  $t_1 = 40$ !



# Regola del modus ponens

Siano date le seguenti proposizioni:

$p =$  Il computer X ha passato con successo il test di Turing

$p \rightarrow q =$  Passare con successo il Test di Turing implica che una macchina può pensare

$q =$  La macchina può pensare.

La regola del “modus ponens” stabilisce che:

$$(p \wedge (p \rightarrow q)) \rightarrow q$$

a parole:

Se passare il Test di Turing implica che la macchina può pensare AND un certo computer ha passato il Test di Turing allora l'implicazione è che la macchina può pensare.

Ovvero:

Se  $p$  è vera e  $p$  implica  $q$ , allora  $q$  è vera.

Infatti:

$P$	$Q$	$P \rightarrow Q$	$P \wedge (P \rightarrow Q)$	$(P \wedge (P \rightarrow Q)) \rightarrow Q$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

In altri termini, è una tautologia (è sempre vera).

# Validità e soddisfacibilità.

Si dice *valida* una fbf con valore vero in *tutte* le interpretazioni.

**Esempio:**

$$P(a) \rightarrow (P(a) \vee P(b))$$

(si ricava dalla tavola di verità che il valore è sempre V).

Se sono presenti quantificatori, le formule diventano infinite: pertanto il calcolo dei predicati è *indecibibile*.

Esistono tuttavia *sottoclassi decidibili* del calcolo dei predicati .

Inoltre si dimostra che se una fbf è valida, esiste una procedura che ne verifica la validità (l'opposto non è vero): pertanto il calcolo dei predicati è *semidecibibile*.

Se una interpretazione dà valore  $V$  a tutte le fbf di un insieme, si dice che le *soddisfa*.

Una fbf  $W$  *segue logicamente* da un insieme di fbf  $S$  se ogni interpretazione che soddisfa  $S$  soddisfa pure  $W$ .

Esempi:

$$(\forall x \forall y) \{P(x) \vee Q(y)\}$$

segue logicamente dall'insieme:

$$\{(\forall x \forall y) \{P(x) \vee Q(y)\}, (\forall z) \{R(z) \vee Q(a)\}\}$$

e così:

$P(a)$  segue logicamente da  $\{(\forall x)P(x)\}$

e ancora

$$(\forall x)Q(x)$$

segue logicamente dall'insieme

$$\{(\forall x) \{\neg P(x) \vee Q(y)\}, (\forall x) P(x)\}$$

*Provare che una fbf  $W$  è conseguenza logica di un dato insieme  $S$  di fbff è riportato alla dimostrazione che  $W$  segue logicamente da  $S$ .*

Nota: per l'indecibilità del calcolo dei predicati, se  $W$  segue da  $S$ , vi è una procedura in grado di dimostrarlo; se  $W$  non segue da  $S$ , tale procedura non sempre è in grado di rilevarlo.



Se  $W$  segue logicamente da  $S$ , una interpretazione che soddisfa  $S$  soddisfa anche  $W$ . Le stesse interpretazioni NON soddisfano  $\neg W$ .

Dunque nessuna interpretazione può soddisfare l'insieme unione di  $S$  e  $\{\neg W\}$  (cioè questo insieme è *insoddisfacibile*).

La chiave della prova sta proprio in questo:  
se  $W$  segue logicamente da  $S$ , l'insieme  $S \cup \{\neg W\}$  è insoddisfacibile e, viceversa, se  $S \cup \{\neg W\}$  è insoddisfacibile,  $W$  segue logicamente da  $S$ .

Per dimostrare che un insieme di fbf è insoddisfacibile, occorre provare che non esiste nessuna interpretazione in cui tutte le fbf hanno valore vero.

Per poter trattare in modo meccanico, le fbf devono essere poste in *forma a clausole*.

# La formula a clausole.

Si applicano in successione alcune operazioni, illustrate mediante il seguente esempio:

$$(\forall x) \{P(x) \rightarrow \{(\forall y) \{P(y) \rightarrow P(f(x,y))\} \wedge \\ \neg(\forall y) \{Q(x,y) \rightarrow P(y)\} \} \}$$

1) Eliminazione delle implicazioni, ovvero

$A \rightarrow B$  è sostituito da:  $\neg A \vee B$

Pertanto:

$$(\forall x) \{ P(x) \rightarrow \{ (\forall y) \{ P(y) \rightarrow P(f(x,y)) \} \wedge \\ \neg(\forall y) \{ Q(x,y) \rightarrow P(y) \} \} \}$$

diventa:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ \neg(\forall y) \{ \neg Q(x,y) \vee P(y) \} \} \}$$

## 2) Riduzione del campo dei segni di negazione (la negazione applicata ad una sola lettera predicativa)

In pratica:

$\neg(A \wedge B)$	è rimpiazzata da	$\neg A \vee \neg B$
$\neg(A \vee B)$	"	$\neg A \wedge \neg B$
$\neg\neg A$	"	$A$
$\neg(\forall x)A$	"	$(\exists x)\{\neg A\}$
$\neg(\exists x)A$	"	$(\forall x)\{\neg A\}$

Pertanto:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ \neg(\forall y) \{ \neg Q(x,y) \vee P(y) \} \} \}$$

diventa:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ (\exists y) \{ \neg \{ \neg Q(x,y) \vee P(y) \} \} \} \}$$

e poi:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ (\exists y) \{ Q(x,y) \wedge \neg P(y) \} \} \}$$

3) Standardizzazione delle variabili: si ribattezzano le variabili quantificate in modo che ogni quantificatore abbia una variabile apparente unica.

In pratica:

$$(\forall x) \{P(x) \rightarrow (\exists x) Q(x)\}$$

diventa:

$$(\forall x) \{P(x) \rightarrow (\exists y) Q(y)\}$$

Pertanto:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ (\exists y) \{ Q(x,y) \wedge \neg P(y) \} \} \}$$

diventa:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ (\exists w) \{ Q(x,w) \wedge \neg P(w) \} \} \}$$



- 4) Eliminazione dei quantificatori esistenziali:  
ogni variabile esistenzialmente quantificata è  
rimpiazzata con una *funzione di Skolem*.

Esempio:

Supponiamo che la fbf

$$(\forall y \exists x)P(x,y)$$

possa essere interpretata come: “per tutti gli  $y$   
esiste un  $x$  tale che  $x$  è maggiore di  $y$ ”.

NB:  $x$  può dipendere da  $y$ !

Allora si cerca una funzione  $g(y)$  (detta funzione di Skolem) che manda ogni valore di  $y$  nell' $x$  che “*esiste*”.

Quindi la fbf diventa

$$(\forall y)P(g(y),y)$$

Si osservi ancora.

- $\exists z$  si elimina in:

$$\{(\forall w)Q(w)\} \rightarrow (\forall x) \{(\forall y) \{(\exists z) \{P(x,y,z) \\ \rightarrow (\forall u)R(x,y,u,z)\} \} \}$$

ottenendo:

$$\{(\forall w)Q(w)\} \rightarrow (\forall x) \{(\forall y) \{P(x,y,g(x,y)) \\ \rightarrow (\forall u)R(x,y,u,g(x,y))\} \}$$

Se il quantificatore esistenziale non si trova nel campo di un quantificatore universale, la funzione di Skolem ha zero argomenti.

Esempio:

$(\exists x)P(x)$  è sostituito da  $P(a)$

dove  $a$  è una costante che sappiamo “esistere”.

L'esempio che stiamo seguendo diventa da così:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ (\exists w) \{ Q(x,w) \wedge \neg P(w) \} \} \}$$

a così:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ \{ Q(x,g(x)) \wedge \neg P(g(x)) \} \} \}$$

dove  $g(x)$  è una funzione di Skolem.

5) Conversione in forma prenessa: tutti i quantificatori universali (che sono tutti diversi) vengono spostati all'inizio della fbf (forma prenessa).

L'esempio da:

$$(\forall x) \{ \neg P(x) \vee \{ (\forall y) \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ \{ Q(x,g(x)) \wedge \neg P(g(x)) \} \} \}$$

diventa:

$$(\forall x \forall y) \{ \neg P(x) \vee \{ \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ \{ Q(x,g(x)) \wedge \neg P(g(x)) \} \} \}$$

dove:

$(\forall x \forall y)$  è detto *prefisso* e

$\{ \neg P(x) \vee \{ \{ \neg P(y) \vee P(f(x,y)) \} \wedge \{ Q(x,g(x)) \wedge \neg P(g(x)) \} \} \}$   
è detta *matrice*

6) Trasformazione della matrice in forma normale congiuntiva: la matrice viene scritta come congiunzione di un numero finito di disgiunzioni di predicati e/o negazioni di predicati (forma normale congiuntiva).

(In parole povere, AND di OR, ovvero prodotti di somme, ovvero, maxterm).



Esempi di forma normale congiuntiva:

$$\{P(x) \vee Q(x,y)\} \wedge \{P(w) \vee \neg R(y)\} \wedge Q(x,y)$$

$$P(x) \vee Q(x,y)$$

$$P(x) \wedge Q(x,y)$$

$$\neg R(y)$$

In pratica si applica ripetutamente la relazione:

$$A \vee (B \wedge C) \equiv \{ A \vee B \} \wedge \{ A \vee C \}$$

L'esempio

$$(\forall x \forall y) \{ \neg P(x) \vee \{ \{ \neg P(y) \vee P(f(x,y)) \} \wedge \\ \{ Q(x,g(x)) \wedge \neg P(g(x)) \} \} \}$$

diventa:

$$(\forall x \forall y) \{ \{ \neg P(x) \vee \neg P(y) \vee P(f(x,y)) \} \wedge \\ \{ \neg P(x) \vee Q(x,g(x)) \} \wedge \{ \neg P(x) \vee \neg P(g(x)) \} \}$$

7) Eliminazione dei quantificatori universali: resta la sola matrice in cui, essendo le variabili legate, sono tutte universalmente quantificate.

8) Eliminazione dei segni di congiunzione.

I segni di congiunzione (cioè  $\wedge$ ; esempio:  $A \wedge B$ ) sono eliminati dando luogo a due fbf (nell'esempio,  $A$  e  $B$ ).

Applicando ripetutamente questo rimpiazzo si ottiene una lista finita di fbf, ognuna delle quali è una disgiunzione ( $\vee$ ) di formule atomiche e/o di negazioni di formule atomiche.

Nomenclatura:

*letterale*: formula atomica (affermata o negata)

*clausola*: fbf costituita solo da disgiunzione di letterali

L'esempio che stiamo seguendo dà:

$\neg P(x) \vee \neg P(y) \vee P(f(x,y))$

$\neg P(x) \vee Q(x,g(x))$

$\neg P(x) \vee \neg P(g(x))$

NB: tutte le variabili si intendono universalmente quantificate

Se in un letterale si sostituiscono al posto delle variabili delle costanti (in generale espressioni prive di variabili) si ottiene un *esempio base*.

Esempio:

dato  $Q(x,y)$ , un esempio base è  $Q(a,f(g(b)))$

## Esempio trasformazione in clausola

“Tutti i romani che conoscono Marco o odiano Cesare o pensano che tutti quelli che odiano qualcuno sono matti”

fbf corrispondente:

$$\begin{aligned} \forall x [\text{romano}(x) \wedge \text{conosce}(x, \text{Marco})] \\ \rightarrow [\text{odia}(x, \text{Cesare}) \vee (\forall y (\exists z \text{odia}(y, z)) \\ \rightarrow \text{credematto}(x, y))] \end{aligned}$$

## 1. eliminazione segni implicazioni:

$$\begin{aligned} \forall x \neg [\text{romano}(x) \wedge \text{conosce}(x, \text{Marco})] \\ \vee [\text{odia}(x, \text{Cesare}) \vee (\forall y \neg (\exists z \text{odia}(y, z)) \vee \\ \text{credematto}(x, y))] \end{aligned}$$

## 2. riduzione portata negazione:

$$\begin{aligned} \forall x [\neg \text{romano}(x) \vee \neg \text{conosce}(x, \text{Marco})] \\ \vee [\text{odia}(x, \text{Cesare}) \vee (\forall y \forall z \neg \text{odia}(y, z) \vee \\ \text{credematto}(x, y))] \end{aligned}$$

3. Standardizzazione variabili: qui nessuna modifica:  
ogni quantificatore lega già una variabile differente

4. Spostamento dei quantificatori:

$\forall x \forall y \forall z$  [eccetera]

5. Eliminazione quantificatori esistenziali:

non ce ne sono



6. Eliminazione prefisso: (resta la matrice)

$$\begin{aligned} & [\neg \text{romano}(x) \vee \neg \text{conosce}(x, \text{Marco})] \\ & \vee [\text{odia}(x, \text{Cesare}) \vee \neg \text{odia}(y, z) \vee \\ & \text{credematto}(x, y)] \end{aligned}$$

7. Trasformazione in congiunzione di disgiunzioni (AND di OR):

$$\begin{aligned} & \neg \text{romano}(x) \vee \neg \text{conosce}(x, \text{Marco}) \\ & \vee \text{odia}(x, \text{Cesare}) \vee \neg \text{odia}(y, z) \vee \\ & \text{credematto}(x, y) \end{aligned}$$

(1 sola clausola)

# Clausole di Horn

Una clausola di Horn è una disgiunzione di letterali in cui *al massimo uno dei letterali è positivo*.

Esempio:

$\neg L \vee \neg B$  e  $\neg L \vee \neg B \vee C$  sono clausole di Horn

$\neg L \vee B \vee C$  non lo è.

## Importanza delle clausole di Horn:

1. Si possono scrivere come implicazioni la cui premessa è una congiunzione di letterali positivi e la cui conclusione è un singolo letterale positivo. Ad esempio,  $\neg L \vee \neg B \vee C$  è equivalente a  $(L \wedge B) \rightarrow C$ .

- Una clausola con *esattamente* un letterale positivo è detta **clausola definita** (sono la base del Prolog). Il letterale positivo è detto **testa**, quelli negativi formano il **corpo** della clausola.
- Una clausola di Horn di un solo letterale positivo è chiamato **fatto**.
- Una clausola di Horn *senza* letterali positivi può essere scritta in forma di implicazione la cui conclusione vale *False*. Ad esempio,  $\neg L \vee \neg B$  è equivalente a  $L \wedge B \rightarrow \text{False}$ . Tali formule sono dette **vincoli di integrità** nel campo dei database.

2. L'inferenza sulle clausole di Horn può essere svolta mediante gli algoritmi di *concatenazione in avanti* e *concatenazione all'indietro* (e i passi di inferenza sono naturali e facili da seguire).
3. Con le clausole di Horn è possibile verificare l'implicazione logica in un tempo che cresce *linearmente* con la dimensione della base di conoscenza.