Teorema di unicità della lettura di una formula

Matteo Acclavio

15/10/2019

Sommario

In questi appunti autocontenuti dimostriamo il teorema di unicità della scrittura ristretto al caso delle formule proposizionali.

I metodi esposti possono essere estesi a qualsiasi insieme di termini di un linguaggio del primo ordine $\mathcal{L}_V$.

1 Sintassi

1.1 Alfabeto e Parole

Un *alfabeto* $\mathcal{A}$ è un insieme di simboli al più numerabile. Una *parola* sull'alfabeto $\mathcal{A}$ è una sequenza di lunghezza finita di simboli in $\mathcal{A}$. La *parola vuota* è la sequenza vuota di simboli[1]. L'insieme delle parole su un alfabeto $\mathcal{A}$ è indicato con $\mathcal{A}^*$.

Per esempio, se consideriamo l'alfabeto $\mathcal{A} = \{A, B\}$, allora le stringhe di caratteri "A", "B", "$AAAB$", "ABA", "$BBBA$" sono parole in $\mathcal{A}^*$, mentre la stringa "C" o la strina infinita "$AAAA\ldots$" non lo sono.

Se $M \in \mathcal{A}^*$ è una parola, definiamo la *lunghezza di* M come il numero $\mathsf{l}[M]$ di simboli che occorrono in M.

Fissato un alfabeto $\mathcal{A}$, diciamo che una parola $N \in \mathcal{A}^*$ è un segmento iniziale di una parola $M \in \mathcal{A}^*$ se $M = NM'$ per una parola $M' \in \mathcal{A}^*$, ovvero se la stringa M è costituita dalla stringa N seguita dalla stringa M'. Similmente, N è un segmento terminale di M se $M = M'N$. Un segmento iniziale (o finale) è proprio se N è una parola non vuota.

1.2 Formule Proposizionali

Consideriamo l'insieme numerabile P i cui elementi chiameremo *variabili proposizionali* e i seguenti simboli

$$\neg \qquad \wedge \qquad \vee \qquad (\qquad)$$

che chiamiamo rispettivamente "not", "e", "o", "parentesi aperta" e "parentesi chiusa".

Per il resto delle note lavoreremo unicamente con l'alfabeto $\mathcal{A} = \mathsf{P} \cup \{\neg, \wedge, \vee, (,)\}$.

[1] A seconda dei testi, la parola vuota può essere indicata con 1 oppure con ϵ.

Definizione 1.1. L'insieme $\mathcal{F}$ delle formule proposizionali costruite su P è il più piccolo sottoinsieme di $\mathcal{A}^*$ tale che:

- $\mathcal{F}$ contiene P;
- se una parola F appartiene a $\mathcal{F}$ allora $\neg F$ appartiene a $\mathcal{F}$;
- ogni volta che le parole F e G appartengono a $\mathcal{F}$ allora anche le parole $(F \wedge G)$ e $(F \vee G)$ appartengono a $\mathcal{F}$.

Detto in altri termini, l'insieme $\mathcal{F}$ è il più piccolo sottoinsieme di $\mathcal{A}^*$ che contiene P e chiuso rispetto alle seguenti operazioni:

- l'operazione unaria
  - *negazione*

$$\begin{array}{rccc} \neg\,\text{-}\, : & \mathcal{A}^* & \rightarrow & \mathcal{A}^* \\ & M & \rightarrow & \neg M \end{array}$$

- le operazioni binarie:
  - *congiunzione*

$$\begin{array}{rccc} \wedge\langle\,\text{-}\,,\,\text{-}\,\rangle : & \mathcal{A}^* \times \mathcal{A}^* & \rightarrow & \mathcal{A}^* \\ & \langle M, N\rangle & \rightarrow & (M \wedge N) \end{array}$$

  - *diusgiunzione*

$$\begin{array}{rccc} \wedge\langle\,\text{-}\,,\,\text{-}\,\rangle : & \mathcal{A}^* \times \mathcal{A}^* & \rightarrow & \mathcal{A}^* \\ & \langle M, N\rangle & \rightarrow & (M \vee N) \end{array}$$

Notazione 1.1. Nello scrivere le operazioni binarie, abbiamo utilizzato le parentesi $\langle$ e $\rangle$ per indicare le parentesi delle coppie fatte da due elementi in $\mathcal{A}^*$. Sebbene questa notazione non sia una notazione standard, per il resto delle note seguiremo questa convensione utilizzando le parentesi $\rangle$ e $\langle$ per gli insiemi ordinati, coppie o successioni. Il motivo di questa scelta è per evitare di fare confusione tra le parentesi) e (sono simboli dell'alfabeto $\mathcal{A}$ e le parentesi $\rangle$ e $\langle$ che utilizziamo nel "linguaggio matematico corrente".

L'insieme $\mathcal{F}$ può essere visto come l'intersezione di tutti gli insiemi che soddisfano certe proprietà, quindi come un fattore comune di questi insiemi. Per questo motivo, questo tipo di definizione è detto "dall'alto" perché si costruisce l'insieme come "raffinamento" di tutti gli insiemi che lo contengono.

È possibile anche dare una definizione alternativa di $\mathcal{F}$ "dal basso", cioè partendo dai suoi elementi atomici e componendoli per costruire tutti is uoi elementi.

Definizione 1.2. Definiamo la successione $\langle \mathcal{F}_n \rangle_{n\in\mathbb{N}}$ di sottoinsiemi di $\mathcal{A}^*$ come segue:

- $\mathcal{F}_0 = \mathsf{P}$;

- $\mathcal{F}_{n+1} = \mathcal{F}_n \cup \{\neg F \mid F \in \mathcal{F}_n\} \cup \{(F\alpha G) \mid F, G \in \mathcal{F}_n,\ \alpha \in \{\wedge, \vee\}\}$

Teorema 1.1.

$$\mathcal{F} = \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$$

Dimostrazione. Possiamo osservare che la successione è crescente, cioè che $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. Quindi che $\mathcal{F}_0 \subset \mathcal{F}_n$ per ogni $n \in \mathbb{N}$ e $\mathcal{F}_0 \subset \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$. Inoltre, se F e G appartengono ad un certo $\mathcal{F}_n$ allora $(F \wedge G)$ e $(F \vee G)$ appartengono a $\mathcal{F}_{n+1}$. Quindi $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ contiene il più piccolo sottoinsieme di $\mathcal{A}^*$ che contiene P e chiuso per le operazioni di negazione, congiunzione, disgiunzione, implicazione e equivalenza.

Per dimostrare l'altra inclusione ci basta far vedere che $\mathcal{F}_n \subset \mathcal{F}$ per ogni $n \in \mathbb{N}$. Per $n = 0$ questa è banalmente vera siccome $\mathcal{F}_0 = \mathsf{P} \subset \mathcal{F}$. Inoltre possiamo procedere per induzione su n. Infatti se $\mathcal{F}_n \subset \mathcal{F}$ allora anche $\mathcal{F}_{n+1} \subset \mathcal{F}$ dato che $\mathcal{F}$ è chiusa rispetto alle operazioni utilizzate per definire $\mathcal{F}_{n+1}$ a partire da $\mathcal{F}_n$. $\square$

Definizione 1.3 (Altezza di una formula)**.** L'altezza $\mathsf{h}[F]$ di una formula $F \in \mathcal{F}$ è definita come il più piccolo intero $n \in \mathbb{N}$ tale che $F \in \mathcal{F}_n$.

Per esempio, se A e B sono variabili proposizionali allora

$$\mathsf{h}[A] = 0 \qquad \mathsf{h}[((A \wedge B) \vee (B \wedge A)) = 2 \qquad \mathsf{h}[\neg\neg\neg\neg(A \vee B)] = 5]$$

Osserviamo che $\mathcal{F}_n$ è l'insieme delle formule di altezza minore o uguale a n e che $\mathcal{F}_{n-1} \setminus \mathcal{F}_n$ è l'insieme delle formule di lunghezza esattamente $n+1$. Inoltre ne segue che per ogni operazione binaria $\alpha \in \{\wedge, \vee\}$ abbiamo

$$\mathsf{h}[\neg A] \leq \mathsf{h}[A] + 1 \qquad\qquad \mathsf{h}[(A\alpha B)] \leq \mathsf{max}\{\mathsf{h}[A], \mathsf{h}[B]\} + 1$$

Sebbene intuitivamente ci si aspetti che le due disequazioni appena scritte siano delle equazioni, questo passaggio richiede la dimostrazione del teorema di unicità della lettura (vedi Corollario 4.8).

2 Dimostrazioni per induzione sull'insieme di formule

Possiamo quindi dimostrare delle proprietà sulle formule con dei ragionamenti per ricorrenza sull'altezza delle formule dimostrando prima per le le formule atomiche (cioè le formule in $\mathsf{P} = \mathcal{F}_0$) e poi mostrando che se questa proprietà è vera per ogni formula in $\mathcal{F}_n$ allora è vera per le formule in $\mathcal{F}_{n+1}$. Sebbene questo ragionamento sia corretto, un metodo più naturale di dimostrare questa proprietà è di dimostrarla per tutte le formule atomiche e poi dimostrare che questa proprietà è *stabile* o che si *preserva* rispetto alle operazioni utilizzate per definire l'insieme a partire dai suoi elementi atomici. In questo modo può dimostrare una proprietà sulle formule senza l'utilizzo dei numeri naturali (necessari per poter effettuare una dimostrazione per recursione).

Esempio 2.1. Vogliamo verificare che $\mathsf{h}[F] < \mathsf{l}[F]$ per ogni $F \in \mathcal{F}$. Se F è una variabile proposizionale allora $\mathsf{h}[F] = 0 < 1 = \mathsf{l}[F]$. Altrimenti $F = \neg F'$ e $\mathsf{h}[\neg F'] \leq \mathsf{h}[F'] + 1 < \mathsf{l}[F'] + 1 = \mathsf{l}[\neg F']$, oppure $F = (G\alpha H)$ con $\alpha \in \{\wedge, \vee\}$ e $\mathsf{h}[(G\alpha H)] \leq \mathsf{max}\{\mathsf{h}[G], \mathsf{h}[H]\} + 1 < \mathsf{max}\{\mathsf{l}[G], \mathsf{l}[H]\} + 1 \leq \mathsf{l}[G] + \mathsf{l}[H] + 1 < \mathsf{l}[G] + \mathsf{l}[H] + 3 = \mathsf{l}[(G\alpha H)]$.

Osservazione 2.1. Questo ci permette di dire una volta per tutte che la parola vuota non è una formula.

Più in generale, presa una proprietà $\mathcal{Y}(M)$ definita sulle parole $M \in \mathcal{A}^*$, possiamo dire

Lemma 2.1. Se $\mathcal{Y}(M)$ una proprietà delle parole $M \in \mathcal{A}^*$ tale che

- $\mathcal{Y}(M)$ è vera per tutte le parole $M \in \mathsf{P}$;
- per ogni parola M e N in $\mathcal{A}^*$ se $\mathcal{Y}(M)$ e $\mathcal{Y}(N)$ sono vere, allora sono vere $\mathcal{Y}(\neg M)$, $\mathcal{Y}((M \wedge N))$ e $\mathcal{Y}((M \vee N))$;

allora $\mathcal{Y}$ è vera per ogni formula F.

Dimostrazione. Definiamo $\mathcal{Z} = \{M \in \mathcal{A}^* \mid \mathcal{Y}(M)\}$. Per definizione $\mathsf{P} \subset \mathcal{Z}$ e $\mathcal{Z}$ è stabile per le operazioni di negazione, congiunzione e disgiunzione. Ne segue che $\mathcal{F} \subset \mathcal{Z}$. $\square$

In modo analogo, abbiamo il seguente

Lemma 2.2. Se $\mathcal{X}(F)$ una proprietà formule $F \in \mathcal{F}$ tale che

- $\mathcal{X}(F)$ è vera per tutte le formule $F \in \mathsf{P}$;
- per ogni formula F e G, se $\mathcal{X}(F)$ e $\mathcal{X}(G)$ sono vere, allora sono vere $\mathcal{X}(\neg F)$, $\mathcal{X}((F \wedge G))$ e $\mathcal{X}((F \vee G))$;

allora $\mathcal{X}$ è vera per ogni formula F.

Dimostrazione. Consideriamo la proprietà sulle parole in $\mathcal{A}^*$ definita nel modo seguente:

$$\mathcal{Y} := \{M \in \mathcal{F} \text{ e } \mathcal{X}(M)\}$$

Si vede immediatamente che se $\mathcal{X}$ soddisfa le ipotesi, allora $\mathcal{Y}$ soddisfa le ipotesi del Lemma 2.1, quindi che $\mathcal{Y}$ è vera per un insieme di parole $\mathcal{Z} \supset \mathcal{F}$. Siccome se parola M soddisfa $\mathcal{Y}$ allora $M \in \mathcal{F}$, allora ne segue che $\mathcal{X}$ è vera per ogni formula in $\mathcal{F}$. $\square$

3 Albero di decomposizione di una formula

Se consideriamo la seguente parola in $\mathcal{A}^*$

$$(((A \wedge (\neg B \vee \neg A)) \wedge (\neg B \vee \neg C)) \vee (C \vee \neg A))$$

non appare immediato se questa parola sia o meno una formula. Per verificare questa affermazione procederemo scomponendo la parola in parole più semplici, utilizzando le operazioni inverse di quelle date per costruire le formule, fino ad ottenere solo formule atomiche che dovrenno essere variabili proposizionali[2]. Quindi definiamo

$$M_1 = ((A \wedge (\neg B \vee \neg A)) \wedge (\neg B \vee \neg C)) \text{ e } M_2 = (C \vee \neg A)$$

da cui ne segue che $M = (M_1 \vee M_2)$. Inoltre definiamo

$$M_{11} = (A \wedge (\neg B \vee \neg A)), \qquad M_{12} = (\neg B \vee \neg C),$$

$$M_{21} = C, \qquad M_{22} = \neg A$$

sapendo che $M_1 = (M_{11} \wedge M_{12})$ e $M_2 = (M_{21} \vee M_{22})$. Continuando questa (de)costruzione definiamo

$$M_{111} = A, \qquad M_{112} = (\neg B \vee \neg A), \qquad M_{1121} = \neg B, \qquad M_{1122} = \neg A,$$

$$M_{11211} = B, \qquad M_{11221} = A$$

$$M_{121} = \neg B, \qquad M_{122} = \neg C, \qquad M_{1211} = B, \qquad M_{1221} = C$$

$$M_{121} = \neg B, \qquad M_{1211} = B, \qquad M_{122} = \neg C, \qquad M_{1221} = C$$

$$M_{221} = A$$

per i quali sappiamo

$$M_{11} = (M_{111} \wedge M_{112}), \qquad M_{112} = (M_{1121} \vee M_{1122}),$$

$$M_{1121} = \neg M_{11211}, \qquad M_{1122} = \neg M_{11221}$$

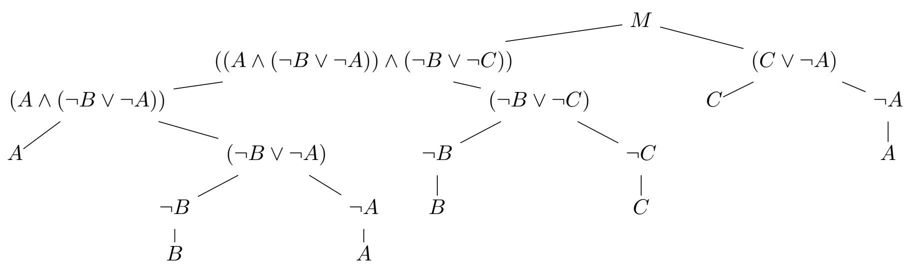
$$M_{12} = (M_{121} \vee M_{122}), \qquad M_{121} = \neg M_{1211}, \qquad M_{122} = \neg M_{1221}$$

$$M_{22} = \neg M_{221}$$

Questo dimostra che M è una formula siccome è ottenuta a partire dalle variabili proposizionali $A, B, C \in \mathsf{P}$ e le operazioni di negazione, congiunzione, disgiunzione e implicazione.

Ad ogni modo, questo modo di rappresentare la decomposizione della parola M fa sanguinare gli occhi. Possiamo rappresentare la decomposizione di questa formula sotto forma di albero:

[2]Questo tipo di tecnica che prevede di decomporre un oggetto nelle sue componenti atomiche al fine di verificarne alcune sue proprietà viene detta *co-induzione*. Questo tipo di tecniche sono ben note in computer science. Il nome fa riferimento al fatto che queste tenciche siano in un qualche modo il duale dell'induzione dove invece un oggetto viene definito a partire da oggetti atomici e delle procedure costruttive.



La radice dell'albero è la formula $M = (((A \wedge (\neg B \vee \neg A)) \wedge (\neg B \vee \neg C)) \vee (C \vee \neg A))$, le foglie sono le variabili proposizionali che occorrono in M e i nodi interni sono le varie sottoformule di M.

Possiamo quindi associare ad ogni parola $M \in \mathcal{A}^*$ un albero $\mathcal{T}_M$ costruito nel modo seguente:

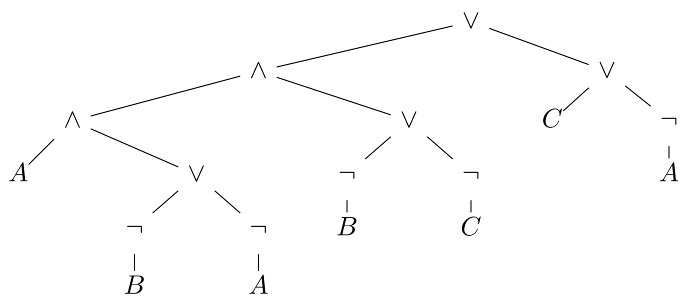
- la radice dell'albero è la parola M;
- le foglie dell'albero sono le variabili proposizionali che occorrono in M;
- ogni nodo $\neg N$ ha un solo figlio N;
- ogni nodo $(N_1 \alpha N_2)$ con $\alpha \in \{\neg, \wedge, \vee\}$ ha due figli N_1 e N_2;
- ogni altro nodo non ha parenti.

Proposizione 3.1. Una parola $M \in \mathcal{A}^*$ è una formula se e solamente ogni foglia dell'albero $\mathcal{T}_M$ è una variabile proposizionale in P.

Dimostrazione. Se $M \in \mathcal{F}$ allora $M \in \mathsf{P} = \mathcal{F}_0$ oppure $M \in \mathcal{F}_{n+1}$. Se $M \in \mathsf{P}$, allora $\mathcal{T}_M$ ha un unico nodo che è radice ma anche foglia dell'albero. Se $M \in \mathcal{F} \setminus \mathsf{P}$ allora $M = \neg N$ per una formula $N \in \mathcal{F}_n$, oppure $M = (N_1 \alpha N_2)$ per due formule $N_1, N_2 \in \mathcal{F}_n$ con $\alpha \in \{\neg, \wedge, \vee\}$.

Viceversa, possiamo procedere per induzione sull'altezza dell'albero $\mathcal{T}_M$. Se l'albero ha altezza 0 allora è costituito dalla sola radice che è anche una foglia, quindi $M \in \mathsf{P}$. Se l'albero ha altezza $n+1$ allora avrà al più due figli e tutti i suoi figli possono essere considerati come radici di un albero di altezza n. Se ha un unico figlio N, allora $M = \neg N$. Analogamente, se ha due figli N_1 e N_2 allora $M = (N_1 \alpha N_2)$ per un $\alpha \in \{\neg, \wedge, \vee\}$. In entrambi i casi, $\mathcal{T}_N$ e $\mathcal{T}_{N_1}$ e $\mathcal{T}_{N_2}$ sono radici di due alberi di altezza strettamente minore dell'altezza di $\mathcal{T}_M$. Concludiamo utilizzando l'ipotesi induttiva. $\square$

Possiamo inoltre rappresentare un albero della formula M evitando la ridondanza di informazioni:



Questa rappresentazione arborescente ci dice che la formula $M = (((A \wedge (\neg B \vee \neg A)) \wedge (\neg B \vee \neg C)) \vee (C \vee \neg A))$ appartiene all'insieme $\mathcal{F}_5$ e quindi che $\mathsf{h}[M] \leq 5$. Ad ora, non abbiamo nessuno strumento che ci permetta di dire che $\mathsf{h}[M] = 5$. Infatti potrebbe esistere un altro albero di decomposizione di M con altezza minore. L'impossibilità di poter decomporre una fomula in due modi differenti verrà dimostrata nella prossima sezione.

4 Unicità della lettura

Per ogni parola $M \in \mathcal{A}^*$, definiamo $\mathsf{o}[M]$ come il numero di parentesi aperte in M, ovvero come il numero di occorrenze del simbolo "(" nella stringa M. In modo analogo definiamo $\mathsf{c}[M]$ come il numero di parentesi chiuse in M.

Lemma 4.1. In ogni formula il numero di parentesi chiuse è uguale a quelle aperte

Dimostrazione. Utilizzando il Lemma 2.2, basta osservare che:

- Se F è una formula atomica $F \in \mathsf{P}$ e $\mathsf{o}[F] = \mathsf{c}[F] = 0$;
- Se $F = \neg G$ allora $\mathsf{o}[F] = \mathsf{o}[G]$ e $\mathsf{c}[F] = \mathsf{c}[G]$. Ne segue che se G è una formula, e quindi $\mathsf{o}[G] = \mathsf{c}[G]$, allora $\mathsf{o}[F] = \mathsf{c}[F]$;
- Se $F = (G\alpha H)$ allora $\mathsf{o}[F] = \mathsf{o}[G] + \mathsf{o}[H] + 1$ e $\mathsf{c}[F] = \mathsf{c}[G] + \mathsf{c}[H] + 1$. Ne segue che se G e H sono una formule, e quindi $\mathsf{o}[G] = \mathsf{c}[G]$ e $\mathsf{o}[H] = \mathsf{c}[H]$, allora $\mathsf{o}[F] = \mathsf{o}[G] + \mathsf{o}[H] + 1 = \mathsf{c}[G] + \mathsf{c}[H] + 1 = \mathsf{c}[F]$;

$\square$

Lemma 4.2. Per ogni formula $F \in \mathcal{F}$ e per ogni parola $M \in \mathcal{A}^*$, se M è un segmento iniziale di F allora $\mathsf{o}[M] \geq \mathsf{c}[M]$.

Dimostrazione. Utilizzando il Lemma 2.2, ci basta osservare che:

- Se F è una formula atomica $F \in \mathsf{P}$ allora l'unico segmento iniziale M di F è F stesso o la parola vuota. Siccome $\mathsf{o}[F] = \mathsf{c}[F] = 0$, in ogni caso ne segue che $\mathsf{o}[M] = \mathsf{c}[M] = 0$;

- Se $F = \neg G$, allora un segmento iniziale M di F è la parola vuota oppure della forma $\neg N$ dove N è un segmento iniziale di G. Nel primo caso abbiamo $\mathsf{o}[M] = \mathsf{c}[M] = 0$. Altrimenti abbiamo che $\mathsf{o}[\neg N] = \mathsf{o}[N]$ e $\mathsf{c}[\neg N] = \mathsf{c}[N]$. Possiamo quindi conludere che $\mathsf{o}[M] = \mathsf{o}[N] \geq \mathsf{c}[N] = \mathsf{c}[M]$ siccome N è un segmento iniziale di G.

- Se $F = (G\alpha H)$, allora abbiamo quattro casi possibili:
  - M è la parola vuota. Allora $\mathsf{o}[M] = \mathsf{c}[M] = 0$;
  - $M = (N$ con N segmento iniziale di una formula G. Per ipotesi N è segmento inziale di una formula e quindi $\mathsf{o}[N] \geq \mathsf{c}[N]$. Ne segue che $\mathsf{o}[M] = 1 + \mathsf{o}[N] \geq 1 + \mathsf{c}[N]$.
  - $M = (G\alpha N$ con N segmento iniziale di una formula H. In modo analogo al caso precedente, per ipotesi N è segmento iniziale di una formula e quindi $\mathsf{o}[N] \geq \mathsf{c}[N]$. Ne segue che $\mathsf{o}[M] = 1 + \mathsf{o}[N] \geq 1 + \mathsf{c}[N]$.
  - $M = F$ allora $\mathsf{o}[M] = \mathsf{o}[G] + \mathsf{o}[H] + 1 = \mathsf{c}[G] + \mathsf{c}[H] + 2 = \mathsf{c}[M]$.

□

Possiamo usare questi due primi risultati per dimostrare che

Lemma 4.3. Per ogni formula $F \in \mathcal{F}$ che inizia con il simbolo della parentesi aperta e per ogni parola $M \in \mathcal{A}^*$, se M è un segmento iniziale proprio di di F allora $\mathsf{o}[F] > \mathsf{c}[F]$.

Dimostrazione. Se una formula F inizia con il simbolo della parentesi aperta, allora $F = (G\alpha H)$ con $\alpha \in \{\wedge, \vee\}$. Siccome M è un segmento iniziale propio, abbiamo solo due casi possibili: $M = (N$ con N segmento iniziale di una formula G oppure $M = (G\alpha N$ con N segmento iniziale di una formula H. Nel primo caso abbiamo che $\mathsf{o}[M] = \mathsf{o}[N] + 1$ e $\mathsf{c}[M] = \mathsf{c}[N]$ mentre nel secondo $\mathsf{o}[M] = 1+\mathsf{o}[G]+\mathsf{o}[N]$ e $\mathsf{c}[M] = \mathsf{c}[G]+\mathsf{c}[N]$. In entrambi ci casi, utilizzando il Lemma 4.2, possiamo concludere che $\mathsf{o}[M] > \mathsf{c}[M]$. Quindi possiamo concludere utilizzando il Lemma 4.1 che siccome $\mathsf{o}[M] \neq \mathsf{c}[M]$ allora M non può essere una formula. □

Lemma 4.4. Per ogni formula $F \in \mathcal{F}$ e per ogni parola $M \in \mathcal{A}^*$, se M è un segmento iniziale proprio di F allora non è una formula.

Dimostrazione. Anche qui utilizziamo il Lemma 2.2:

- Se F è una formula atomica, allora F non ha segmenti iniziali propri;

- Se $F = \neg G$ allora ogni suo segmento iniziale proprio sarà della forma $\neg$ oppure $\neg N$ con N segmento iniziale proprio di G. Per definizione la parola $\neg$ non è una formula. Inoltre, ipotesi induttiva è un segmento iniziale proprio di una formula G e quindi N non è una formula.

  Per concludere ci basta dimostrare che $\neg N$ non è una formula. Infatti, se lo fosse, allora dovrebbe esistere una formula H tale che $\neg N = \neg H$. Questo implicherebbe che $N = H$ e che quindi N è una formula contraddicendo l'ipotesi induttiva.

- Se $F = (G\alpha H)$ e M un segmento proprio di F allora possiamo utilizzare il Lemma 4.3 per concludere M non può essere una formula dato che $\mathsf{o}[M] > \mathsf{c}[M]$.

$\square$

Abbiamo quindi ora tutti gli strumenti per dimostrare il seguente.

Teorema 4.5 (Lettura unica di una formula proposizionale)**.** Se F è una formula allora solo una delle seguenti affermazioni è vera:

- F è una variabile proposizionale in P;
- esiste una formula G tale che $F = \neg G$;
- esiste un unico connettivo binario α e un'unica coppia di formule G e H in $\mathcal{F}$ tale che $F = (G\alpha H)$.

Dimostrazione. I tre casi si escludono a vicenda. Infatti una formula può iniziare con una variabile proposizionale, il simbolo $\neg$ oppure il simbolo della parentesi aperta.

- se $F \in \mathsf{P}$ e quindi decomposizione è unica visto che la formula non può essere decomposta;
- se F inizia con il simbolo $\neg$ allora abbiamo che $F = \neg G$. Se esistesse un secondo G' tale che $F = \neg G'$ allora avremmo che $G = G'$;
- se F inizia con il simbolo della parentesi aperta allora $F = (G\alpha H)$. Se esistesse un secondo modo di scrivere $F = (K\beta L)$ allora avremmo un'equivalenza delle parole $G\alpha H$ e $K\beta L$. Questo implicherebbe, senza perdere di generalità, che K è segmento iniziale di G. Per il Lemma 4.4 K non può essere segmento iniziale proprio. Siccome la parola vuota non è un segmento iniziale proprio, e quindi $K = G$. Ciò implica l'ugualianza delle due parole αH e βL.

$\square$

Ne seguono i seguenti corollari

Corollario 4.6 (Unicità albero di decomposizione)**.** Se F è una formula in $\mathcal{F}$ allora il suo albero di decomposizione è unico.

Corollario 4.7. Se F è una formula in $\mathcal{F}$, allora $\mathsf{o}[F] = \mathsf{c}[F]$.

Corollario 4.8. Se F è una formula in $\mathcal{F}$, allora $\mathsf{h}[F]$ è uguale alla lunghezza massima di un ramo dell'albero $\mathcal{T}_F$.