Sommario

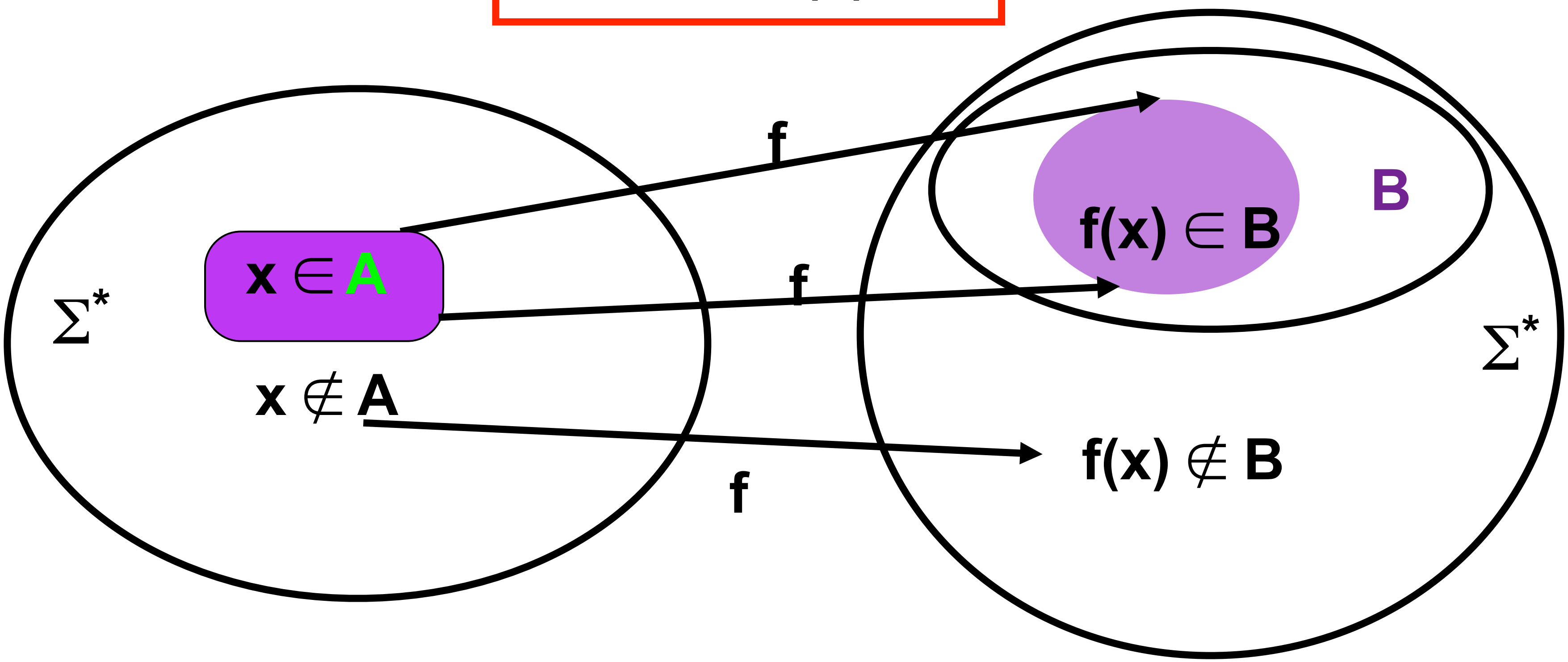
- Riduzioni polinomiali (alla Karp)
- NP- completezza
- Proprietà dei problemi NP- completi
- Teorema di Cook

Riduzione polinomiale

Dati A $\subseteq \Sigma^*$ e B $\subseteq \Sigma^*$ due linguaggi diciamo che A <u>si riduce polinomialmente a</u> B, in breve A $\leq_p$ B, se esiste una funzione

f : $\Sigma^* \rightarrow \Sigma^*$ <u>Turing calcolabile</u> in tempo polinomiale tale che

$$\boxed{x \in A \Leftrightarrow f(x) \in B}$$

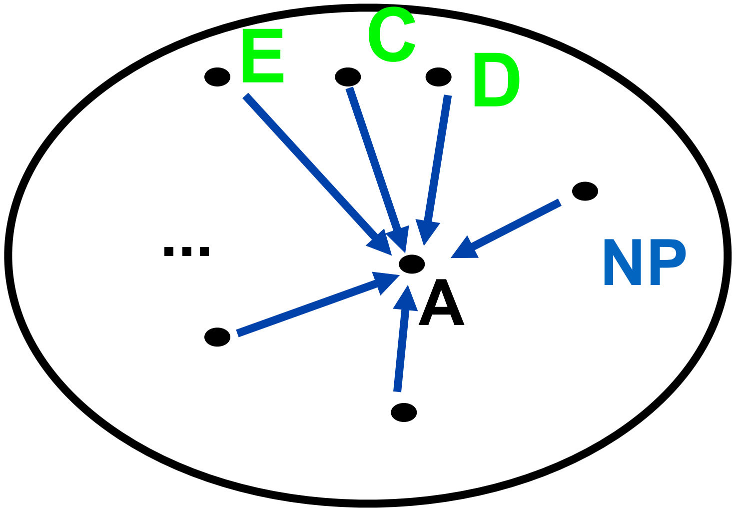


NP-completezza

Un problema A (linguaggio A) si dice NP-completo se

1. **A $\in$ NP**
2. **è NP-hard cioè è tale che ogni problema in NP si riduce polinomialmente ad esso**

   ($\forall$ B $\in$ NP B $\leq_p$ A)



Riduzioni polinomiali: proprietà

Teorema 1. Se B $\leq_p$ A e A $\in$ P allora B $\in$ P.

Teorema 2. La relazione $\leq_p$ è riflessiva e transitiva.

Teorema 3. Se A $\in$ P, A $\in$ NP-completo allora P = NP

Teorema 4.
Se A $\in$ NP-completo , A $\leq_p$ B con B $\in$ NP allora B $\in$ NP-completo

Formule booleane

- **Una formula booleana della logica proposizionale è induttivamente definita come ottenuta a partire da un insieme numerabile di variabili booleane, i connettivi $\wedge$, "e", $\vee$,"o",$\neg$"non"e le parentesi.**

- **Una formula è soddisfacibile se esiste un assegnamento di valori di verità alle variabili che rende la formula vera.**

Esempi: $(X \wedge Y) \vee (\neg X \wedge Z)$ è soddisfatta da $X=Y=Z=1$, mentre $X \wedge \neg X$ non può essere soddisfatta

$(x \vee \neg y \vee z) \wedge (\neg x \vee y)$ è una CNF formula

NP-completezza: teorema di COOK

Sia SAT il linguaggio delle formule soddisfacibili e sia CNFSAT il linguaggio delle formule soddisfacibili in CNF.

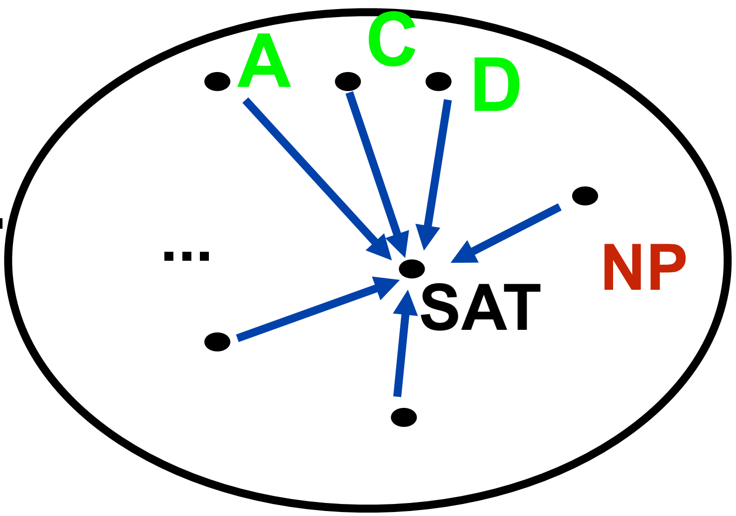
Teorema di COOK: SAT (e CNFSAT) è NP-completo

Teorema di COOK: inizio prova

Per dimostrarlo si deve far vedere che

1. **SAT è in NP**
2. **ogni linguaggio A in NP si riduce polinomialmente a SAT.**

Il punto 1 già visto.
Dimostreremo il punto 2.



Teorema di COOK: la prova

Se A è in NP esiste una NTM $M = (Q,\Sigma,\Gamma,\delta,q_0,q_a,q_r)$ che decide A con complessità di tempo n^k, per qualche k. Utilizzando M e il polinomio n^k definiremo una funzione f, polinomialmente Turing-calcolabile, che associa una formula φ ad ogni x, in modo tale che $x \in A$ sse $\varphi \in$ SAT.

La formula φ deve descrivere una sequenza di configurazioni

determinata dall'input x: $c_0 \Rightarrow_T c_1 \Rightarrow_T \dots \Rightarrow_T c_i$

con $c_0 = q_0 x$ e c_i di accettazione, $c_i = \alpha q_a \beta$, se $x \in L(M)$, con $1 \leq i \leq n^k$

Serviranno quindi delle formule per esprimere proprietà che assicurino che ogni c_j è una configurazione (un solo simbolo per cella di nastro, un solo stato, una sola posizione per la testina di lettura), e proprietà riguardanti la relazione "porta a" tra configurazioni:

se $(p,b,R) \in \delta(q,a)$ allora $\alpha q a c \beta \Rightarrow_T \alpha b p c \beta$

se $(p,b,L) \in \delta(q,c)$ allora $\alpha a q c \beta \Rightarrow_T \alpha p a b \beta$

Teorema di COOK: le variabili della formula

Data la NTM $M = (Q,\Sigma,\Gamma,\delta,q_0,q_a,q_r)$ che decide A con complessità di tempo n^k, per qualche k, la formula associata φ utilizza le seguenti variabili:

$S_{i,q}$ per ogni $i=1,...,n^k$ e $q \in Q$ che dice, se vera, "q è lo stato di M all'i-simo passo"

$T_{i,j,a}$ per ogni $i,j=1,...,n^k$ e $a \in \Gamma$ che dice, se vera, "a è il simbolo in lettura nella j-sima posizione sul nastro di M all'i-simo passo"

$H_{i,j}$ per ogni $i,j=1,...,n^k$ che dice, se vera, "la testina di lettura è nella j-sima posizione sul nastro di M all'i-simo passo"

Teorema di COOK: la formula 1° parte

La formula φ è la congiunzione delle seguenti :

➔ φ_{start} **che dice, se vera, "all'inizio, cioè al passo 1 del calcolo, lo stato è q_0, la testina di lettura è sulla prima cella del nastro che contiene $x = a_1,\dots,a_n$ seguito da n^k - n occorrenze di $\perp$ "**

➔ $\varphi_{start} = S_{1,q_0} \wedge H_{1,1} \bigwedge_{1\leq j\leq n} T_{1,j,a_j} \bigwedge_{n<j\leq n^k} T_{1,j,\perp}$

dove $\bigwedge_{1\leq i\leq n^k} X_{i,j} = X_{1,j} \wedge \dots \wedge X_{n^k,j}$

Teorema di COOK: la formula 1° parte

➛ $\varphi_{testina}$ che dice, se vera, " la <u>testina di lettura</u> è posizionata su un'unica cella del nastro in ogni passo di calcolo"

$$\varphi_{testina} = \bigwedge_{1 \le i \le n^k} \varphi_{testina,i}$$

➛ $\varphi_{testina,i}$ che dice, se vera, " la <u>testina di lettura</u> è posizionata su un'unica cella del nastro al passo i-simo di calcolo"

$$\varphi_{testina,i} = \left(\bigvee_{1 \le j \le n^k} H_{i,j} \right) \wedge \left(\bigwedge_{1 \le j < j' \le n^k} \neg \left(H_{i,j} \wedge H_{i,j'} \right) \right)$$

ma non su due!

se vera, la testina è su una cella

la cui equivalente in CNF

$$\varphi_{testina,i} = \left(\bigvee_{1 \le j \le n^k} H_{i,j} \right) \wedge \left(\bigwedge_{1 \le j < j' \le n^k} \left(\neg H_{i,j} \vee \neg H_{i,j'} \right) \right)$$

Faremo vedere in ogni passo anche come costruire la formula φ in CNF.

Teorema di COOK: la formula 2° parte

L' analoga a $\varphi_{testina}$ per simboli di nastro :

- φ_{nastro} che dice, se vera, “ogni cella del nastro contiene un solo simbolo in ogni passo di calcolo”

$\varphi_{nastro} = \bigwedge_{1 \leq i \leq n^k} \bigwedge_{1 \leq j \leq n^k} \varphi_{nastro,i,j}$ e $\varphi_{nastro,i,j}$ dice, se vera, “la cella j-sima del nastro contiene un solo simbolo al passo i-simo di calcolo”

$$\varphi_{nastro,i,j} = \left(\bigvee_{\alpha \in \Gamma} T_{i,j,\alpha}\right) \wedge \left(\bigwedge_{\alpha \neq \alpha' \in \Gamma} \neg\left(T_{i,j,\alpha} \wedge T_{i,j,\alpha'}\right)\right)$$

l'equivalente in CNF:

$$\varphi_{nastro,i,j} = \left(\bigvee_{\alpha \in \Gamma} T_{i,j,\alpha}\right) \wedge \left(\bigwedge_{\alpha \neq \alpha' \in \Gamma} \left(\neg T_{i,j,\alpha} \vee \neg T_{i,j,\alpha'}\right)\right)$$

Teorema di COOK: la formula 2° parte

L' analoga a $\varphi_{testina}$ per simboli di stato:

➔ φ_{stato} che dice, se vera, " M si trova in un unico stato in ogni passo di calcolo"

$\varphi_{stato} = \bigwedge_{1 \leq i \leq n^k} \varphi_{stato,i}$ e $\varphi_{stato,i}$ dice, se vera, "M si trova in un unico stato al passo i-simo di calcolo"

$$\varphi_{stato,i} = \left(\bigvee_{q \in Q} S_{i,q}\right) \wedge \left(\bigwedge_{q \neq q' \in Q} \neg(S_{i,q} \wedge S_{i,q'})\right)$$

con la sua equivalente in CNF:

$$\varphi_{stato,i} = \left(\bigvee_{q \in Q} S_{i,q}\right) \wedge \left(\bigwedge_{q \neq q' \in Q} (\neg S_{i,q} \vee \neg S_{i,q'})\right)$$

Teorema di COOK: la formula 3° parte

→ $\varphi_{mossa} = \bigwedge_{1\le i\le n^k} \bigwedge_{1\le j\le n^k} \bigwedge_{\alpha\in\Gamma} \bigwedge_{q\in Q} \varphi_{i,j,\delta(q,\alpha)}$ che dice, se vera, "ogni passo di calcolo è conforme a δ"

$\varphi_{i,j,\delta(q,\alpha)}$ dice, se vera, "l'i-simo passo di calcolo, con la testina di lettura sulla cella j-sima, è conforme a δ"

Sia $\delta(q,\alpha) = \{(q_1,\alpha_1,D_1),\ldots,(q_m,\alpha_m,D_m)\}$, con $m\ge 1$

$$\varphi_{i,j,\delta(q,\alpha)} = (H_{i,j} \wedge S_{i,q} \wedge T_{i,j,\alpha}) \Rightarrow$$

$$((H_{i+1,j_1} \wedge S_{i+1,q_1} \wedge T_{i+1,j,\alpha_1}) \vee \ldots \vee (H_{i+1,j_m} \wedge S_{i+1,q_m} \wedge T_{i+1,j,\alpha_m}))$$

dove se $D_k = L$ allora la nuova posizione della testina di lettura $j_k = j-1$, se $j>1$ e $j_k=1$ se $j=1$, mentre se $D_k = R$ allora $j_k = j+1$, per $1\le k\le m$.

La sua versione equivalente in CNF:

$$\bigwedge_{1\le i\le n^k} \bigwedge_{1\le j\le n^k} \bigwedge_{\alpha\in\Gamma} \bigwedge_{q\in Q} \Big((\neg H_{i,j} \vee \neg S_{i,q} \vee \neg T_{i,j,\alpha} \vee H_{i+1,j_1}) \wedge$$

$$(\neg H_{i,j} \vee \neg S_{i,q} \vee \neg T_{i,j,\alpha} \vee S_{i+1,q_1}) \wedge (\neg H_{i,j} \vee \neg S_{i,q} \vee \neg T_{i,j,\alpha} \vee T_{i+1,j,\alpha_1}) \wedge$$

$$\ldots \wedge (\neg H_{i,j} \vee \neg S_{i,q} \vee \neg T_{i,j,\alpha} \vee T_{i+1,j,\alpha_m}) \Big)$$

Teorema di COOK: la formula 4° parte

➔ $\varphi_{mossa/nastro}$ che dice, se vera, "solo il <u>simbolo in lettura</u> può cambiare in un passo di calcolo"

➔ $\varphi_{mossa/nastro} = \bigwedge_{1\leq i\leq n^k} \bigwedge_{1\leq j\leq n^k} \bigwedge_{1\leq j\neq j'\leq n^k} \bigwedge_{\alpha\in\Gamma} (H_{i,j} \wedge T_{i,j',\alpha}) \Rightarrow T_{i+1,j',\alpha}$

con la sua versione in CNF:

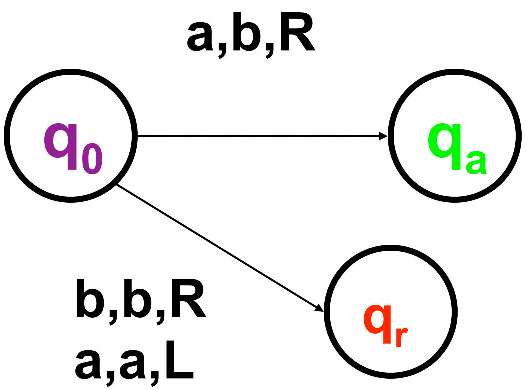
$\bigwedge_{1\leq i\leq n^k} \bigwedge_{1\leq j\leq n^k} \bigwedge_{1\leq j\neq j'\leq n^k} \bigwedge_{\alpha\in\Gamma} (\neg H_{i,j} \vee \neg T_{i,j',\alpha} \vee T_{i+1,j',\alpha})$

➔ $\varphi_{accettazione}$ che dice, se vera, "al passo i-simo del calcolo M entra nello <u>stato di accettazione</u> "

➔ $\varphi_{accettazione} = \bigvee_{1\leq i\leq n^k} (S_{i,q_a})$

Teorema di COOK: esempio di φ -1

$$\varphi = \varphi_{start} \land \varphi_{testina} \land \varphi_{nastro} \land \varphi_{stato} \land \varphi_{mossa} \land \varphi_{mossa/nastro} \land \varphi_{accettazione}$$



1 1 1 1

$$\varphi_{start} = S_{1,q_0} \land H_{1,1} \land T_{1,1,a} \land T_{1,2,\perp}$$

$$\varphi_{testina} = \varphi_{testina,1} \land \varphi_{testina,2}$$

0

$$\varphi_{testina,1} = (H_{1,1} \lor H_{1,2}) \land \neg (H_{1,1} \land H_{i,2})$$

$$\varphi_{testina,2} = (H_{2,1} \lor H_{2,2}) \land \neg (H_{2,1} \land H_{2,2})$$

0 1

$$\varphi_{nastro} = \varphi_{nastro,1,1} \land \varphi_{nastro,1,2} \land \varphi_{nastro,2,1} \land \varphi_{nastro,2,2}$$

0 0

$$\varphi_{nastro,1,1} = (T_{1,1,a} \lor T_{1,1,b} \lor T_{1,1,\perp}) \land (\neg(T_{1,1,a} \land T_{1,1,b}) \land \neg (T_{1,1,a} \land T_{1,1,\perp}) \land \neg (T_{1,1,b} \land T_{1,1,\perp}))$$

1

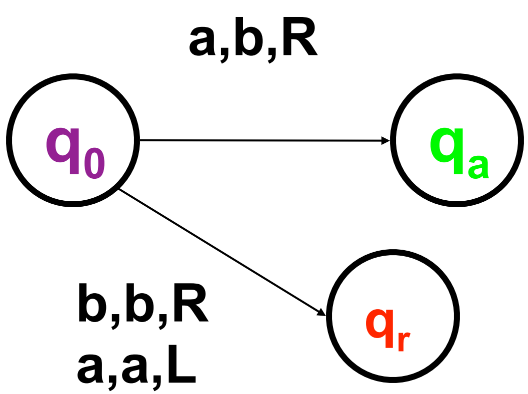
0 0

$$\varphi_{stato} = \varphi_{stato,1} \land \varphi_{stato,2}$$

$$\varphi_{stato,1} = (S_{1,q_0} \lor S_{1,q_a} \lor S_{1,q_r}) \land (\neg(S_{1,q_0} \land S_{1,q_a}) \land \neg(S_{1,q_0} \land S_{1,q_r}) \land \neg(S_{1,q_a} \land S_{1,q_r})$$

Teorema di COOK: esempio di φ -2

$$\varphi = \varphi_{start} \wedge \varphi_{testina} \wedge \varphi_{nastro} \wedge \varphi_{stato} \wedge \varphi_{mossa} \wedge \varphi_{mossa/nastro} \wedge \varphi_{accettazione}$$



$= \varnothing$

$$\varphi_{mossa} = \varphi_{1,1,\delta(q_0,a)} \wedge \varphi_{1,1,\delta(q_0,b)} \wedge \varphi_{1,1,\delta(q_0,\perp)} \wedge$$
$$\varphi_{1,2,\delta(q_0,a)} \wedge \varphi_{1,2,\delta(q_0,b)} \wedge \varphi_{1,2,\delta(q_0,\perp)} \wedge$$
$$\varphi_{2,1,\delta(q_0,a)} \wedge \varphi_{2,1,\delta(q_0,b)} \wedge \varphi_{2,1,\delta(q_0,\perp)} \wedge$$
$$\varphi_{2,2,\delta(q_0,a)} \wedge \varphi_{2,2,\delta(q_0,b)} \wedge \varphi_{2,2,\delta(q_0,\perp)} \wedge \ldots$$

$$\varphi_{1,1,\delta(q_0,a)} = ((\overset{1}{H_{1,1}} \wedge \overset{1}{S_{1,q_0}} \wedge \overset{1}{T_{1,1,a}}) \Rightarrow (\overset{1}{H_{2,2}} \wedge \overset{1}{S_{2,q_a}} \wedge \overset{1}{T_{2,1,b}}) \vee (\overset{0}{H_{2,1}} \wedge \overset{0}{S_{2,q_r}} \wedge \overset{0}{T_{2,1,a}}))$$

$$\varphi_{1,1,\delta(q_0,b)} = ((\overset{1}{H_{1,1}} \wedge \overset{1}{S_{1,q_0}} \wedge \overset{0}{T_{1,1,b}}) \Rightarrow (\overset{0}{H_{2,2}} \wedge \overset{0}{S_{2,q_r}} \wedge \overset{1}{T_{2,1,b}}))$$

$$\varphi_{mossa/nastro} = (\overset{1}{H_{1,1}} \wedge \overset{1}{T_{1,2,\perp}}) \Rightarrow \overset{1}{T_{2,2,\perp}} \wedge (H_{1,1} \wedge T_{1,2,b}) \Rightarrow T_{2,2,b} \wedge$$
$$(H_{1,1} \wedge T_{1,2,a}) \Rightarrow T_{2,2,a} \wedge (H_{1,2} \wedge T_{1,1,a}) \Rightarrow T_{2,1,a} \wedge (H_{1,2} \wedge T_{1,1,b}) \Rightarrow T_{2,1,b} \ldots$$

$$\varphi_{accettazione} = (S_{1,q_a} \vee S_{2,q_a})$$

Teorema di COOK: complessità

Abbiamo visto che x $\in$ A sse $\varphi \in$ SAT, perchè se x è in A allora ha un cammino di accettazione e quindi possiamo costruire un assegnamento di verità per φ, d'altro canto se $\varphi \in$ SAT all'assegnamento di valori di verità corrisponde un cammino di accettazione per x.

La formula φ ha lunghezza $O(n^{2k})$ perchè il numero degli stati o dei simboli di nastro sono costanti, mentre il numero delle occorrenze delle variabili è $O(n^{2k})$. Guardiamo per esempio:

$$\varphi_{mossa} = \bigwedge_{1 \le i \le n^k} \bigwedge_{1 \le j \le n^k} \bigwedge_{\alpha \in \Gamma} \bigwedge_{q \in Q} \varphi_{i,j,\delta(q,\alpha)}$$

notiamo che il numero di letterali è $O(n^{2k})$.

NP e NP-completi: lo scenario più probabile