DIMOSTRAZIONE DEL TEOREMA DI FAGIN - PRIMA PARTE

TEOREMA. $\exists SO = NP$

$\exists SO = \{ struct(F) \mid F$ È FORMULA AL SECOND'ORDINE ESISTENZIALE $\}$

$\exists SO \subseteq NP$

- IN QUEST'INCLUSIONE CI OCCUPIAMO DI DIMOSTRARE CHE OGNI FORMULA F ESISTENZIALE AL SECOND'ORDINE È TALE PER CUI ESISTE UNA MdT NONDETERMINISTICA E POLYTIME M_F CHE DECIDE PROPRIO $struct(F)$
- ABBIAMO BISOGNO DI UN PAIO DI LEMMI AUSILIARI

LEMMA 1

- OGNIQUALVOLTA ESISTA ALMENO UN SIMBOLO PREDICATIVO DI ARIETÀ ALMENO PARI AD 1, VALE CHE

$$|bin^n(I)| \geq n$$

DIMOSTRIAMOLO

- SE ESISTE COME PER IPOTESI, UN SIMBOLO PREDICATIVO P_j DI ARIETÀ ALMENO PARI AD 1, ALLORA AVREMO CHE

$$|bin^n(I)| = |bin^n(P_1) \dots bin^n(P_m) \; bin^n(f_1) \dots bin^n(f_k)|$$

$$\geq |bin^n(P_j)| = n^{ar(P_j)} \geq n$$

$\boxtimes$

LEMMA 2

$FO \subseteq P$

DIMOSTRIAMOLO

- DIMOSTRARE $FO \subseteq P$ SIGNIFICA DIMOSTRARE CHE PER OGNI F CHIUSA NELLA LOGICA AL PRIMORDINE, $struct(F) \subseteq P$
- NON POSSIAMO PROCEDERE QUINDI PER INDUZIONE PERCHÉ F POTREBBE AVERE SOTTOFORMULE APERTE, ALLE QUALI NON SI PUÒ APPLICARE L'IPOTESI INDUTTIVA.
- OCCORRE QUINDI DIMOSTRARE UN RISULTATO LEGGERMENTE PIÙ FORTE, OVVERO IL SEGUENTE:

  PER OGNI F CON VARIABILI LIBERE $x_1, \dots, x_m$ ESISTE UN ALGORITMO $\mathcal{A}_F$ POLYTIME TALE CHE SU INPUT $S, i_1, \dots, i_m$ DETERMINA SE $S = bin^n(I)$ DOVE

$$(A_n, I), \xi \models F$$

DOVE $\xi(x_j) = i_j$

QUESTO È EFFETTIVAMENTE UNO STATEMENT CHE POSSIAMO DIMOSTRARE PER INDUZIONE SULLA STRUTTURA DI F:

- SE $F: P(t_1, \dots, t_p)$ ALLORA A_F PROCEDERÀ NEL MODO SEGUENTE:
  - PRIMA DI TUTTO CALCOLANDO $[\![t_i]\!]_\xi$ DOVE ξ È L'AMBIENTE CHE ASSEGNA i_j AD x_j
  - POI, CONTROLLA CHE L'INTERPRETAZIONE DI P, RICAVABILE DA S SIA TALE PER CUI $([\![t_1]\!]_\xi, \dots [\![t_p]\!]_\xi)$ APPARTIENE A TALE INTERPRETAZIONE.

  OSSERVIAMO CHE IN QUESTO MODO A_F DETERMINA CORRETTAMENTE SE

$$(A_n, I), \xi \models F.$$

- SE $F = F_1 \wedge F_2$, ALLORA A_F LO COSTRUIRÒ A PARTIRE DA A_{F_1} E A_{F_2}, I QUALI ESISTONO PER IPOTESI INDUTTIVA. IN PARTICOLARE A_F RITORNERÀ IL VALORE 1 SSE A_{F_1} E A_{F_2} RITORNANO IL VALORE 1
- SE $F = F_1 \vee F_2$ O $F = \neg F_1$, ALLORA

SI PROCEDE ESATTAMENTE COME NEL CASO PRECEDENTE

- SE $F = \exists x. G$ ALLORA PROCEDIAMO USANDO L'IPOTESI INDUTTIVA E IL LEMMA 1. PER L'IPOTESI INDUTTIVA, INFATTI, A_G ESISTE POLYTIME. INOLTRE A_G SI ASPETTA ANCHE UN INPUT i_q RELATIVO PROPRIO ALLA VARIABILE x. CIÒ CHE FARÀ A_F È CHIAMARE A_G PIÙ VOLTE, UNA PER OGNI VALORE POSSIBILE DI i_q. POICHÉ IL NUMERO DI TALI VALORI POSSIBILI È n E PER IL LEMMA 1, $|bin^n(I)| \geq n$ A_F PRENDERÀ TEMPO POLINOMIALE. IL RISULTATO RESTITUITO DA A_F SARÀ INFINE 1 SSE A_G RITORNA 1 ALMENO UNA VOLTA.

- SE $F = \forall x. G$, ALLORA POSSIAMO PROCEDERE ANALOGAMENTE AL CASO PRECEDENTE

$\boxtimes$

$\boxed{\exists SO \subseteq NP}$

$FO \subseteq P$ $\rightsquigarrow$ LA VOLTA SCORSA

$\exists SO \subseteq NP$

DIMOSTRIAMOLO

- RICORDIAMO INNANZITUTTO CHE UNA FORMULA "DI" $\exists SO$ È NELLA FORMA

$$G \equiv \exists X_1^{n_1} \dots \exists X_m^{n_m} . F$$

DOVE F È UNA FORMULA AL PRIM'ORDINE.

- POSSIAMO DIRE CHE IL PROBLEMA DI VERIFICARE SE $bin^n(I) \in struct(G)$ PUÒ ESSERE RISOLTO CONTROLLANDO CHE $bin^n(J) \in struct(F)$, DOVE J È UN'INTERPRETAZIONE CHE ESTENDE I CON DELLE STRINGHE CHE INTERPRETANO $X_1^{n_1}, \dots X_m^{n_m}$.
ABBIAMO INFATTI CHE

$$(\mathcal{A}_n, I) \models G$$

$$\Updownarrow$$

ESISTE J TALE CHE J ESTENDE I

$$(\mathcal{A}_n, J) \models F\{R_1/X_1^{n_1}, \dots, R_m/X_m^{n_m}\}$$

DOVE $R_1, \dots, R_m$ SONO SIMBOLI CHE <u>NON</u> OCCORRONO IN F E J INTERPRETA TALI SIMBOLI IN UN MODO QUALUNQUE

- A QUESTO PUNTO È CHIARO COME POSSA ESSERE COSTRUITO UN ALGORITMO DI DECISIONE NONDETERMINISTICO E POLYTIME PER $struct(G)$:

1. ESTRAE DALLA STRINGA IN INPUT IL PARAMETRO n.

2. GENERA m STRINGHE BINARIE IN MODO NONDETERMINISTICO, CIASCUNA CORRISPONDENTE AD UNA POSSIBILE INTERPRETAZIONE PER R_i
3. MODIFICHERÀ LA STRINGA IN INPUT USANDO LE STRINGHE COSTRUITE AL PUNTO 2 IN MODO DA FAR DIVENTARE LA PRIMA UNA CODIFICA DI UNA CERTA INTERPRETAZIONE J PER $F\{R_1/X_1^{n_1}, \dots, R_m/X_m^{n_m}\}$.
4. CHIAMEREMO POI L'ALGORITMO POLYTIME PER LA DECISIONE DI

$$struct\left(F\{R_i/X_i^{n_i}\}\right)$$

CHE SAPPIAMO COSTRUIRE GRAZIE AL LEMMA $FO \subseteq P$.

OSSERVIAMO COME L'ALGORITMO CHE ABBIAMO COSTRUITO SIA CORRETTO, NONDETERMINISTICO E POLYTIME. DI CONSEGUENZA, $struct(G) \in NP$

$\boxtimes$

$\boxed{NP \subseteq \exists SO}$

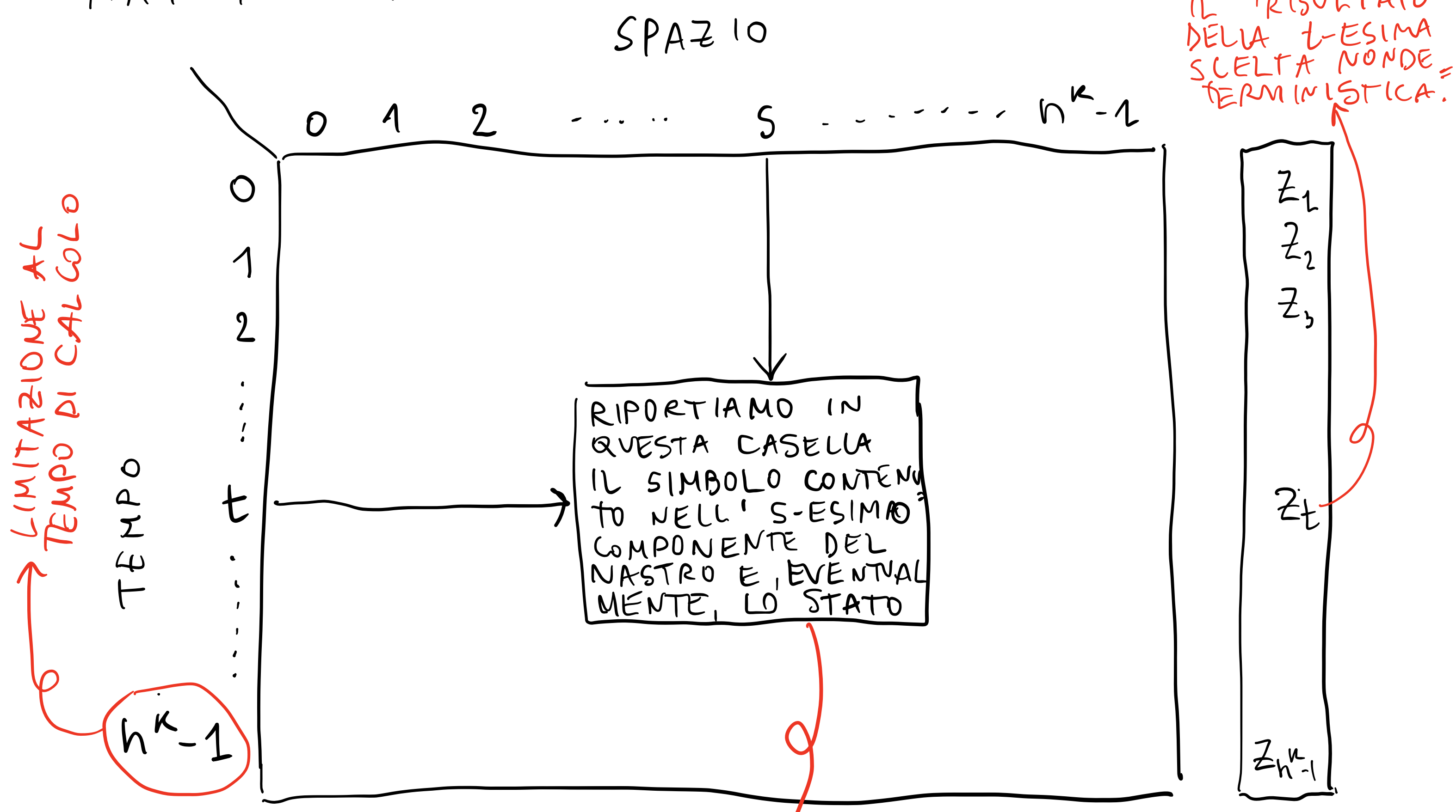
- SI TRATTA ORA DI CAPIRE SE <u>OGNI</u> PROBLEMA IN NP POSSA DIVENTARE UNA FORMULA "DI" $\exists SO$.
- PROCEDEREMO FACENDO VEDERE CHE OGNI MACCHINA NONDETERMINISTICA E POLYTIME M CORRISPONDE ED È CODIFICABILE IN UNA FORMULA F_M IN MODO TALE CHE

$$struct(F_M) = \{x \in \{0,1\}^* \mid M \text{ ACCETTA } x\}$$

- F_M SARÀ UNA FORMULA CHE USA, OLTRE ALLE VARIABILI AL SECOND'ORDINE, <u>UN</u> SINGOLO

SIMBOLO PREDICATIVO UNARIO, CHE INDICHIAMO CON S.

- LA FORMULA F_M "VEDE" L'ESECUZIONE DI M SU UN CERTO INPUT COME UNA MATRICE AVENTE LA FORMA SEGUENTE



GLI ELEMENTI DI QUESTA MATRICE SARANNO ELEMENTI DELL'INSIEME

$$\Sigma \uplus (\Sigma \times Q)$$

IL SOLO SIMBOLO, QUANDO LA TESTINA NON E' "LI'"

SIMBOLO E STATO, QUANDO LA TESTINA È PROPRIO "LI'"

- LA FORMULA CHE VOGLIAMO COSTRUIRE AVRÀ LA FORMA SEGUENTE

$$F_M = \exists C_1^{2k} \ldots \exists C_g^{2k} . \exists \Delta^k . \varphi_M$$

DOVE:

- g È $|\Sigma_1 \uplus (\Sigma_1' \times Q)|$, OVVERO IL NUMERO DI VALORI DIVERSI CHE POSSIAMO TROVARE IN UNA CASELLA DELLA MATRICE
- k È L'ESPONENTE DEL POLINOMIO $n^k - 1$
- IL PREDICATO C_i^{2k} (DOVE $i \in \{1, \dots, g\}$) È UN PREDICATO CHE INDICA, INTUITIVAMENTE SE LA CASELLA DI INDICE (t, s) È EFFETTIVAMENTE i, QUEST'ULTIMO CODIFICA DEL CORRISPONDENTE VALORE DI $\Sigma_1 \uplus (\Sigma_1 \times Q)$. AD ESEMPIO

$$\underbrace{C_2^{2K}}_{\substack{\text{STA PER IL}\\ \text{"SECONDO"}\\ \text{ELEMENTO}\\ \text{DI}\\ \Sigma_1 \uplus (\Sigma_1 \times Q)}} \Big(\overbrace{\underbrace{t_1, \dots, t_k}_{\substack{\text{UNA CODIFICA}\\ \text{DELL'INDICE DI}\\ \text{RIGA. INFATTI}\\ [\![t_1]\!] \in \{1, \dots, n\}\\ \text{E QUINDI I}\\ \text{POSSIBILI VALORI}\\ \text{DI } [\![t_1]\!], \dots, [\![t_k]\!]\\ \text{SARANNO } n^k}}}^{k \text{ TERMINI}}, \overbrace{\underbrace{s_1, \dots, s_k}_{\substack{\text{UNA CODIFICA}\\ \text{DELL'INDICE DI}\\ \text{COLONNA}}}}^{k \text{ TERMINI}} \Big)$$

VALE 1 SE NELLA CASELLA DI COORDINATE $([\![t_1]\!], \dots, [\![t_k]\!])$ E $([\![s_1]\!], \dots, [\![s_k]\!])$ C'È PROPRIO IL SIMBOLO CORRISPONDENTE A 2, OSSIA IL SECONDO ELEMENTO DI $\Sigma_1 \uplus (\Sigma_1 \times Q)$.

- Δ^k AVRÀ IL RUOLO DI CATTURARE LE SCELTE NONDETERMINISTICHE, NEL MODO SEGUENTE:

$\Delta^k(t_1, \dots, t_k)$

VARRÀ 0 SE ALL'ISTANTE $([\![t_1]\!], \dots, [\![t_k]\!])$ LA MACCHINA M SEGUE LA PRIMA DELLE DUE STRADE NONDETERMINISTICHE, 1 ALTRIMENTI

- LA FORMULA φ_M SARÀ IN REALTÀ UNA FORMULA CHE PRESCRIVE COME I VARI VALORI DI $C_1, \dots C_g$ SONO TRA LORO LEGATI, COME RIFLETTONO L'INPUT E COME INFLUENZANO L'OUTPUT. PIÙ NELLO SPECIFICO

$$\varphi_M = \alpha_M \wedge \beta_M \wedge \gamma_M \wedge \delta_M$$

$C_i(\bar{0}, \bar{s})$ CODIFICA L'INPUT

$C_i(\bar{t}, \bar{s})$ E $C_j(\bar{t}, \bar{s})$ NON SONO IN CONTRADDIZIONE OVVERO AL PIÙ UNO DEI DUE VALE

CODIFICA DELLA FUNZIONE DI TRANSIZIONE DI M

ALL'ISTANTE $n^k - 1$ LA MACCHINA SI TROVA IN UNO STATO DI ACCETTAZIONE

VEDIAMO UN PO' PIÙ NELLO SPECIFICO COME SIA POSSIBILE DEFINIRE QUESTE FORMULE

$$\alpha_M = \forall x . \left(\neg S(x) \rightarrow C^{2k}_{\langle 0 \rangle} (\underbrace{0, \dots, 0}_{k \text{ VOLTE}}, \underbrace{0, \dots, 0}_{k-1}, x) \right)$$

LA CODIFICA DI $0 \in \Sigma$ IN $\{1, \dots g\}$

$$\wedge \left(S(x) \rightarrow C^{2k}_{\langle 1 \rangle} (\underbrace{0, \dots, 0}_{k \text{ VOLTE}}, \underbrace{0 \dots 0}_{k-1}, x) \right)$$

$$\left(\forall y_2 \dots y_k \ . \ C^{2k}_{\langle BLANK \rangle}(\underbrace{0,\dots,0}_{k \text{ VOLTE}},1,y_2\dots y_k)\right)$$

$$\wedge \left(\forall y_1, y_3, \dots y_k \ . \ C^{2k}_{\langle BLANK \rangle}(0,\dots,0,y_1,1,y_3\dots y_k)\right)$$

$$\vdots$$

$$\beta_M = \forall \bar{x} . \forall \bar{y} . \bigwedge_{i \neq j} C_i(\bar{x},\bar{y}) \to \neg C_j(\bar{x},\bar{y})$$

γ_M È IMPOSSIBILE A DESCRIVERSI IN MODO ESPLICTO. SE, AD ESEMPIO LA FUNZIONE DI TRANSIZIONE DI M DICESSE CHE

$$(q,a) \longmapsto (q',a',\leftarrow),(q'',a'',\rightarrow)$$

ALLORA SCRIVEREI QUALCOSA COME

$$\forall \bar{x} . \forall \bar{y} . \left[C^{2k}_{\langle q,a \rangle}(\bar{x},\bar{y}) \longrightarrow \right.$$
$$(\Delta(x) \longrightarrow C^{2k}_{\langle q',a' \rangle}(\bar{x}+1,\bar{y}+1)) \wedge$$
$$\left.(\neg\Delta(x) \longrightarrow C^{2k}_{\langle q'',a'' \rangle}(\bar{x}+1,\bar{y}-1))\right]$$

OCCORRE ANCHE SPECIFICARE CHE TUTTE LE ALTRE POSIZIONI SUL NASTRO, ALL'ISTANTE $\bar{x}+1$ RIMANGONO INVARIATE

COSA SIGNIFICA QUESTA COSA? OCCORRE CATTURARE L'ADDIZIONE E LA SOTTRAZIONE DI UN ELEMENTO ALL'INTERNO DELLA LOGICA

δ_M È MOLTO SEMPLICE. BASTA

STIPULARE CHE LO STATO DELLA MACCHINA ALL'ISTANTE $n^k - 1$ SIA UNO STATO DI ACCETTAZIONE q_{ACC}.

$$\bigvee_{d \in \Sigma} . \exists y_1 \ldots \exists y_n . \; C^{2k}_{\langle q_{ACC}, d \rangle}(\overline{max}, \bar{y})$$

LA CORRETTEZZA DI QUESTA CODIFICA LA POSSIAMO DIMOSTRARE, SENZA ANDARE NEI DETTAGLI, NEL MODO SEGUENTE: SE I CODIFICA LA STRINGA $r \in \{0,1\}^*$ ALLORA PER OGNI AMBIENTE ξ, SE

PER INDUZIONE SU t.

$$(\mathcal{A}_n, I), \xi \models \alpha_M \wedge \beta_M \wedge \gamma_M \wedge \delta_M$$

ALLORA $\xi(C^{2k}_i)$ AVRÀ IL SUO VALORE CORRETTO, OVVERO $\xi(C^{2k}_i)$ VARRÀ IN CORRISPONDENZA DELL'ISTANTE t-ESIMO E DELLA CASELLA s-ESIMA <u>SSE</u> LA MACCHINA M SU INPUT r DOPO t PASSI SI TROVA NELLO STATO i IN POSIZIONE s.