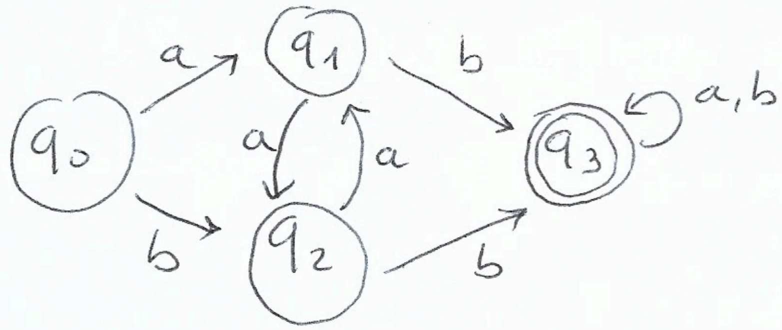
Minimizzazione

N



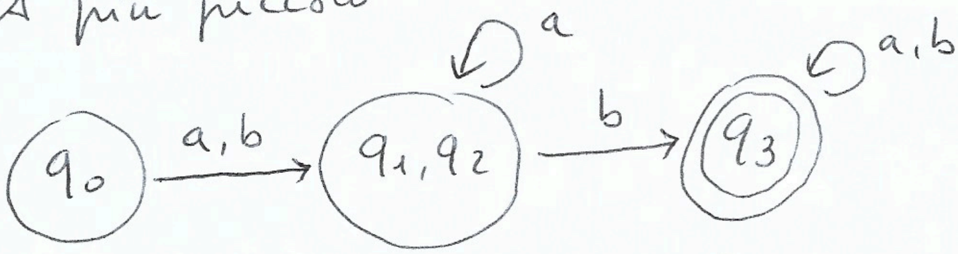
DFA

$$L[N] = L[N, q_0] = (a|b)\, a^*\, b\, (a|b)^*$$

$$L[N, q_1] = a^*\, b\, (a|b)^* = L[N, q_2]$$

$\Rightarrow$ q_1 e q_2 sono due stati equivalenti (o indistinguibili)

Se fondiamo insieme q_1 e q_2, otteniamo un DFA più piccolo



Questo DFA è <u>minimo</u> perché non ci sono due stati tra loro equivalenti. Infatti

q_3 è <u>finale</u> e quindi $L[N, q_3] \ni \varepsilon$

$q_1 q_2$ è <u>non finale</u> e quindi $\varepsilon \notin L[N, q_1 q_2]$
ma $b \in L[N, q_1 q_2]$

q_0 è <u>non finale</u> ma $b \notin L[N, q_0]$

Prima delle definizioni formali, ci serve introdurre una notazione:

Per un DFA $N = (\Sigma, Q, \delta, q_0, F)$

$\hat{\delta} : Q \times \Sigma^* \to Q$ è definita come

$$\hat{\delta}(q, \varepsilon) = q$$
$$\hat{\delta}(q, xa) = \delta(\hat{\delta}(q, x), a)$$

e quindi

$w \in L[N]$ sse $\hat{\delta}(q_0, w) \in F$

Equivalenza / Indistinguibilità

Def Due stati q_1 e q_2 di un DFA N sono equivalenti (o indistinguibili) se $\forall x \in \Sigma^*$

$$\hat{\delta}(q_1, x) \in F \quad \text{sse} \quad \hat{\delta}(q_2, x) \in F$$

cioè se $L[N, q_1] = L[N, q_2]$

Simmetricamente, due stati q_1 e q_2 non sono equivalenti se $\exists x \in \Sigma^*$ tale che

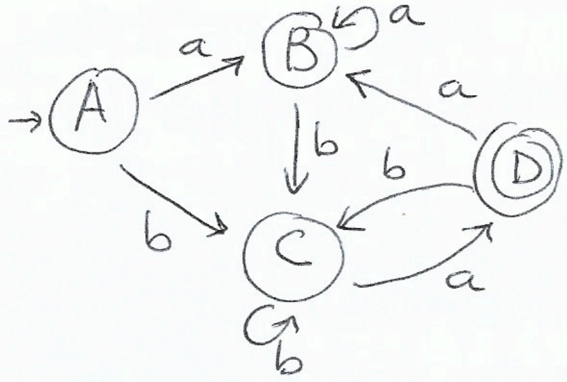
$$\hat{\delta}(q_1, x) \in F \quad \text{ma} \quad \hat{\delta}(q_2, x) \notin F$$

o

$$\hat{\delta}(q_1, x) \notin F \quad \text{ma} \quad \hat{\delta}(q_2, x) \in F$$

q_1 e q_2 sono "distinguibili"

Strategia: cerco di distinguere due stati considerando le $x \in \Sigma^*$ a partire dalla più corta (ε)



Cerco di vedere quali coppie di stati NON sono equivalenti, a cominciare dalla stringa ε

1) ε distingue ogni stato in F da ogni stato in $Q \setminus F$

~~(A, D)~~ ~~(B, D)~~ ~~(C, D)~~

2) Consideriamo ora le stringhe di lunghezza 1, ovvero "a" e "b".

2.1) • "a" distingue B e C perché

$$\delta(B,a) = B \quad e \quad \delta(C,a) = D$$

ma (B, D) l'avevo già cancellata al passo precedente ($B \in Q \setminus F$, mentre $D \in F$)

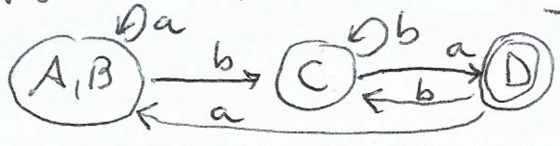
• "a" distingue A e C perché

$$\delta(A,a) = B \quad e \quad \delta(C,a) = D$$

ma $B \in Q \setminus F$ mentre $D \in F$

(2.2) "b" non mi permette di fare ulteriori distinzioni

3) Procedo ora con le stringhe lunghe 2, ma non riesco a fare nessuna ulteriore distinzione $\Rightarrow$ FINE A e B sono equivalenti.

$\Rightarrow$ 

Dato un DFA $M = (\Sigma, Q, \delta, q_0, F)$, definiamo una famiglia di relazioni $\sim_i \subseteq Q \times Q$ nel seguente modo:

- $$\sim_0 \; = \underbrace{F \times F \;\cup\; (Q \setminus F) \times (Q \setminus F)}$$

  stati che non possono essere distinti da ε, l'unica parola di lunghezza 0.

- $q_1 \sim_{i+1} q_2$ sse $\forall a \in \Sigma \quad \delta(q_1, a) \sim_i \delta(q_2, a)$

  che significa: q_1 e q_2 sono in relazione $\sim_{i+1}$ se $\forall x \in \Sigma^*$ con $|x| \leq i+1$

  $$\hat{\delta}(q_1, x) \in F \text{ sse } \hat{\delta}(q_2, x) \in F$$

<u>Oss:</u> (1) La relazione $Id = \{(q,q) \mid q \in Q\}$ è tale che $Id \subseteq \sim_i$ per ogni i

"Uno stato è sempre equivalente a se stesso"

(2) $\sim_i$ è una relazione d'equivalenza per ogni i

- $\sim_0$ ha 2 sole classi d'equivalenza, ovvero F e $Q \setminus F$
- per $\sim_i$, riflessività e simmetria sono ovvie, mentre la transitività è meno banale:

  $q_1 \sim_i q_2$ e $q_2 \sim_i q_3 \Rightarrow q_1 \sim_i q_3$

(3) $\sim_{i+1} \subseteq \sim_i$ Ad ogni passo, rimuovo qualche coppia!

$$\sim_0 \supseteq \sim_1 \supseteq \sim_2 \supseteq \sim_3 \dots$$

catena decrescente di relazioni di equivalenza (non crescente)

(4) Se esiste K tale che $\sim_K = \sim_{K+1}$, allora $\forall J > K$ vale che $\sim_J = \sim_K$ (* vedi prossima pag.)

"cioè, non appena la relazione non viene modificata in un passo, ho trovato "la soluzione""

(5) Un tale K esiste sicuramente ed è minore di $|\sim_0| = |F \times F| + |(Q \setminus F) \times (Q \setminus F)| = |F|^2 + |Q \setminus F|^2$ perché, nella peggiore delle ipotesi, ad ogni passo iterativo rimuovo solo una (meglio 2) coppia.

(In realtà si può dimostrare che $K \leq |Q| - 1$, perché $|Q| - 1$ è la lunghezza del massimo cammino aciclico: infatti, negli esempi che abbiamo visto, ho considerato solo stringhe "corte" che non portavano a cicli) ~~[illegible]~~

Vogliamo dimostrare, con un esempio, che se $\sim_2 = \sim_1$, allora anche $\sim_3 = \sim_2$.

Supponiamo, per assurdo, che $\sim_2 = \sim_1$ ma che $\sim_3 \neq \sim_2$ cioè $\exists A, B$ tali che $A \sim_2 B$ ma $A \not\sim_3 B$, cioè

$$\sim_3 \subset \sim_2 = \sim_1$$

Allora, poiché $A \not\sim_3 B$, deve essere

$$A \xrightarrow{a_1} A_1 \xrightarrow{a_2} A_2 \xrightarrow{a_3} A_3 \qquad A_3 \in F$$
$$B \xrightarrow{a_1} B_1 \xrightarrow{a_2} B_2 \xrightarrow{a_3} B_3 \qquad B_3 \notin F$$

(o viceversa)

Ma allora $A_1 \not\sim_2 B_1$ e poiché $\sim_2 = \sim_1$, deve essere $A_1 \not\sim_1 B_1$. Ma allora

$$A_1 \xrightarrow{b} A_4$$
$$B_1 \xrightarrow{b} B_4 \qquad A_4 \in F,\ B_4 \notin F \text{ (o viceversa)}$$

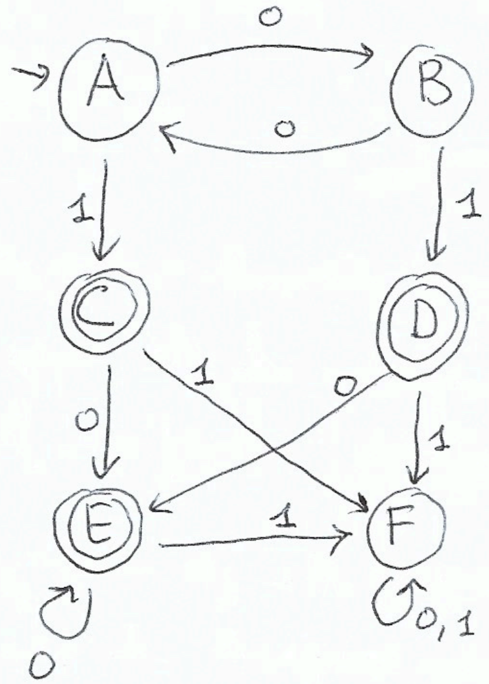
Ma allora $A \not\sim_2 B$ perché

$$A \xrightarrow{a_1} A_1 \xrightarrow{b} A_4$$
$$B \xrightarrow{a_1} B_1 \xrightarrow{b} B_4$$

contraddicendo l'ipotesi iniziale.

$\Rightarrow$ non è possibile che $\sim_1 = \sim_2$ e $\sim_3 \subset \sim_2$!

$$\boxed{\text{Cioè se } \sim_{K+1} = \sim_K, \text{ allora } \forall J > K \ \sim_j = \sim_K}$$



$$\sim_0 = \{(A,A), (A,B), (A,F), (B,B), (B,A), (B,F), (F,F), (F,A), (F,B)\}$$
$$\cup\ \{(C,C), (C,D), (C,E), (D,D), (D,C), (D,E), (E,E), (E,C), (E,D)\}$$

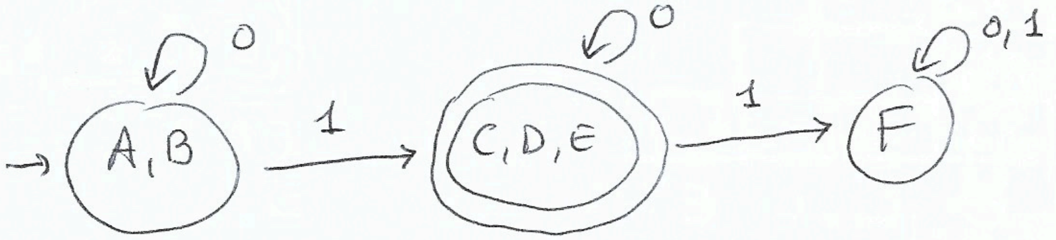
$$\sim_1 = \{(A,A), (A,B), (B,B), (B,A), (F,F), (C,C), (C,D), (C,E), (D,D), (D,C), (D,E), (E,E), (E,C), (E,D)\}$$

$\sim_2 = \sim_1$ OK - FINITO!

Quali sono gli stati equivalenti? Classi d'eq.?

- A e B
- F
- C, D, E

Automa minimo risultante dalla fusione degli stati equivalenti



$$L = 0^* 1 0^*$$

$$q_0 \xrightarrow{a} q_1 \xrightarrow{a} q_2 \xrightarrow{a} \cdots \xrightarrow{a} q_{n-1} \xrightarrow{a} q_n \circlearrowleft a$$

$$\sim_0 \;=\; \{\{q_0, q_1, \ldots q_{n-1}\}, \{q_n\}\}$$

$$\sim_1 \;=\; \{\{q_0, q_1, \ldots, q_{n-2}\}, \{q_{n-1}\}, \{q_n\}\}$$

$$\vdots$$

$$\sim_{n-1} = \{\{q_0\}, \{q_1\}, \{q_2\}, \ldots, \{q_{n-1}\}, \{q_n\}\}$$

= Identità (ogni stato è equivalente solo a se stesso!)

Esempio di caso pessimo: la relazione finale $\sim$ è ottenuta dopo un numero di iterazioni pari al numero n, con $|Q| = n+1$

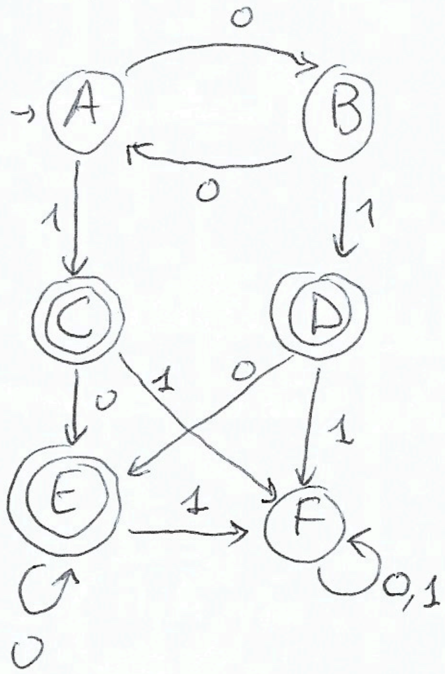
(da 0 a $n-1$)

Possiamo ricavare un algoritmo pratico da questa idea?

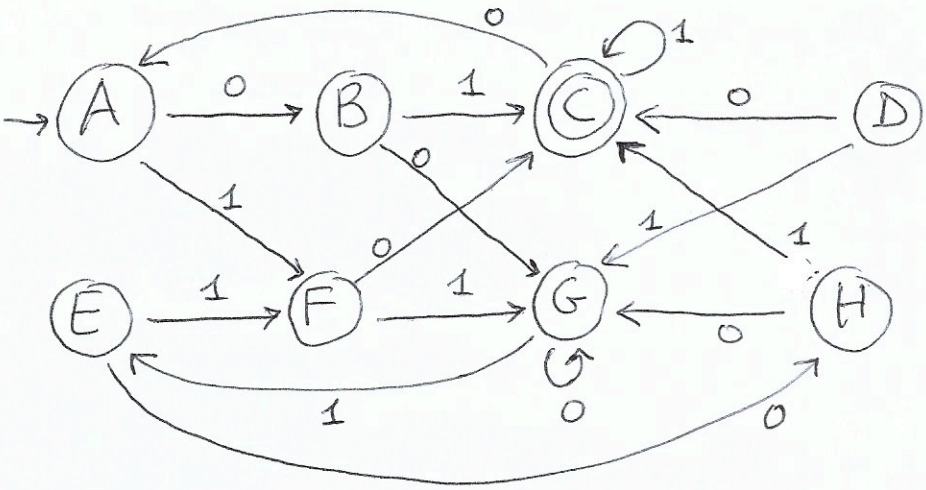
- Tabella con solo coppie "vere"
- al round 0 dell'algoritmo iterativo, metto una marca X_0 per segnalare che la coppia è distinta (finale, nonfinale) (nonfinale, finale)
- al round 1, metto la marca X_1 per distinguere le coppie (q_1, q_2) non ancora marcate che per qualche $a \in \Sigma$ ha $(\delta(q_1, a), \delta(q_2, a))$ già marcata
- al round 2, metto la marca X_2 per ...

se non riesco a mettere nessuna nuova marca in un round $\Rightarrow$ STOP!

	A	B	C	D	E
B					
C	X_0	X_0			
D	X_0	X_0			
E	X_0	X_0			
F	X_1	X_1	X_0	X_0	X_0



Le entrate della tabella che rimangono non marcate mi dicono quali stati sono equivalenti



← dall'ultimo al 2° →

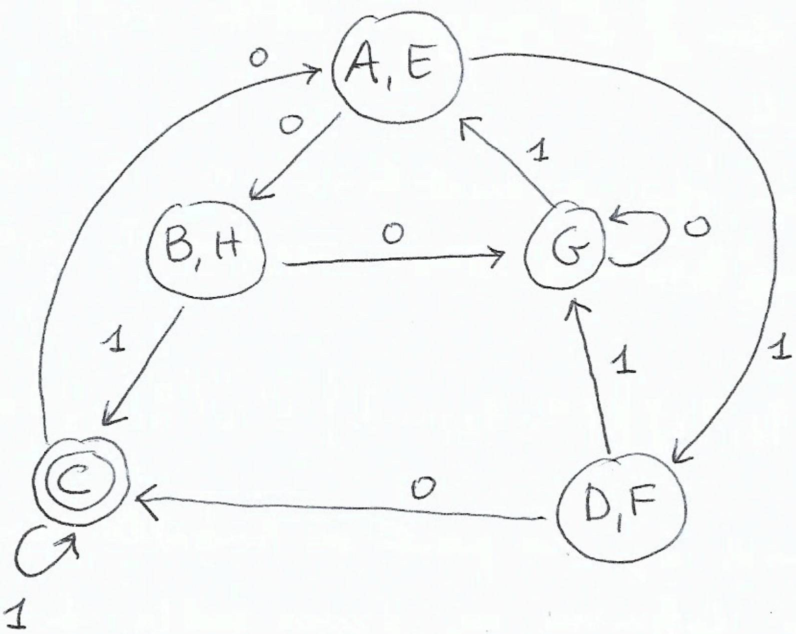
	A	B	C	D	E	F	G
B	X_1						
C	X_0	X_0					
D	X_1	X_1	X_0				
E		X_1	X_0	X_1			
F	X_1	X_1	X_0		X_1		
G	X_2	X_1	X_0	X_1	X_2	X_1	
H	X_1		X_0	X_1	X_1	X_1	X_1

← dal primo al penultimo →

- X_0 finale/non finale
- X_1 $\delta(A,1) = F$
  $\delta(B,1) = C$
  $(C,F)\ X_0$
  ⋮
- X_2 $\delta(A,1) = F$
  $\delta(G,1) = E$
  $(E,F)\ X_1$

  $\delta(E,1) = F$
  $\delta(G,1) = E$
  $(E,F)\ X_1$

$A \sim E$ $\quad B \sim H$ $\quad D \sim F$



Automa Minimo Risultante

Algoritmo Iterativo (Tabella a Scala)

0) Costruire la tabella a scala

1) Marca X_0 ogni coppia (q_1, q_2) tale che $q_1 \in F$ e $q_2 \in Q \setminus F$ (o viceversa)

2) $b := true$; $i := 1$;

3) **while** b **do** {

    3.1) – $b := false$;

    3.2) – per ogni coppia (q_1, q_2) non marcata
    do { **if** $\exists a \in \Sigma$ con $(\delta(q_1, a), \delta(q_2, a))$ già marcata
    then { – marca (q_1, q_2) con X_i;
    – $b := true$ }
    }

    3.3) – $i := i + 1$

}

4) Al termine sia J l'insieme delle coppie non marcate

5) La relazione di equivalenza $\sim$ è la chiusura riflessiva e simmetrica di J, cioè

$$\sim \;= J \cup \{(q_2, q_1) \mid (q_1, q_2) \in J\} \cup \{(q, q) \mid q \in Q\}$$

N.B. al round i, la parte di tabella non marcata definisce $\sim_i$ (una volta chiusa riflessivamente e simmetricamente)

<u>Teorema</u> Dato un DFA $M = (\Sigma, Q, \delta, q_0, F)$, l'algoritmo di riempimento della Tabella a scala termina. Due stati p e q sono distinguibili (ovvero non equivalenti) sse la casella (p,q) (o (q,p)) è marcata (e quindi sono equivalenti sse la casella non è marcata.)

<u>Dim:</u> • Termina sempre: poiché $\exists k.\ \sim_k = \sim_{k+1}$ e quindi l'algoritmo iterativo termina entro k round.

• $\Rightarrow$) Se p e q sono distinguibili, allora $\exists x \in \Sigma^*$. $\hat{\delta}(p,x) \in F$ e $\hat{\delta}(q,x) \notin F$ (o viceversa). Se prendo $k = |x|$, allora di sicuro $(p,q) \notin \sim_k$, cioè (p,q) viene marcata entro il round k

• $\Leftarrow$) Se (p,q) sono marcati, allora sicuramente (p,q) sono distinguibili;

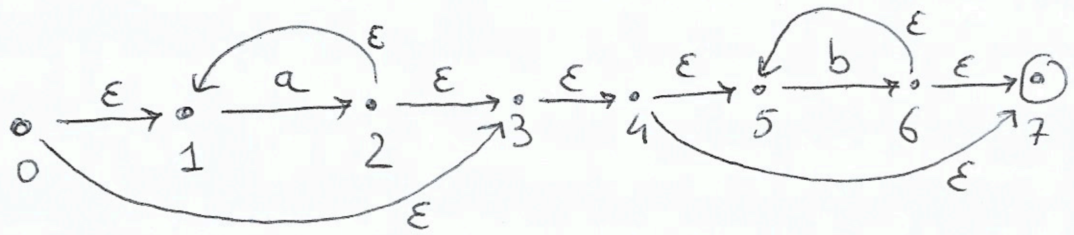
basta prendere la "catena" di coppie/simbolo che portano ad una coppia non presente in $\sim_0$. Ad esempio

$$\underset{X_2}{(p,q)} \xrightarrow{a} \underset{X_1}{(p',q')} \xrightarrow{b} \underset{X_0}{(p'',q'')}$$

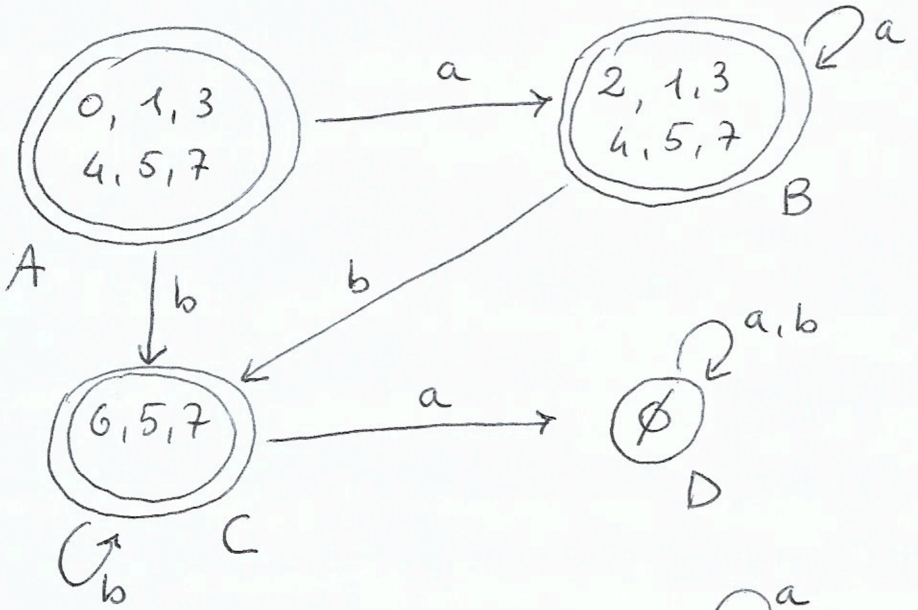
$\Rightarrow$ ab è la stringa che distingue p e q.

Esempio

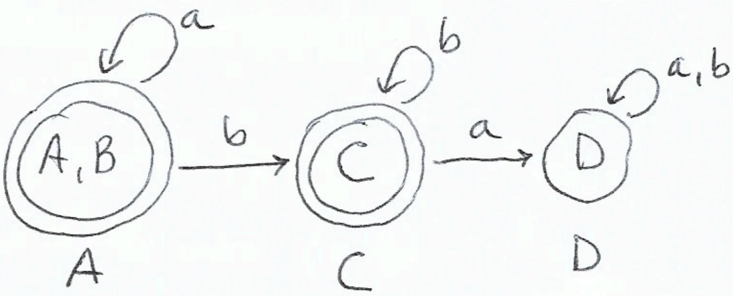
1) $a^* \varepsilon b^*$ Costruire l'NFA associato



2) Ricavare il DFA equivalente



3) Minimizzare il DFA



B			
C	X_1	X_1	
D	X_0	X_0	X_0
	A	B	C

4) Ricavare la grammatica regolare dal DFA minimo

$A \to aA \mid a \mid bC \mid b \mid \varepsilon$ ← perché lo stato è sia iniziale che finale

$C \to bC \mid b \mid aD$

$D \to aD \mid bD$

Alternativamente

$A \to aA \mid \varepsilon \mid bC$

$C \to bC \mid \varepsilon \mid aD$

$D \to aD \mid bD$

5) Semplificare la grammatica, rimuovendo i simboli inutili

$$A \to aA \mid a \mid bC \mid b \mid \varepsilon \qquad A \to aA \mid \varepsilon \mid bC$$
$$C \to bC \mid b \qquad C \to bC \mid \varepsilon$$

6) Ricavare dalla grammatica semplificata, l'espressione regolare associata

$C \simeq b^*b$, e sostituendo

$A \simeq aA \mid a \mid bb^*b \mid b \mid \varepsilon$

cioè

$A \simeq a^*(a \mid bb^*b \mid b \mid \varepsilon)$

$C \simeq b^*$, e sostituendo

$A \simeq aA \mid \varepsilon \mid bb^*$

cioè

$A \simeq a^*(\varepsilon \mid bb^*)$

Nota che l'espressione regolare finale è diversa da quella di partenza, ma è ad essa equivalente!
Si può dimostrare, ad esempio, che

$$a^* \varepsilon b^* \simeq a^*(\varepsilon \mid bb^*)$$

Questo esempio dimostra che questo diagramma

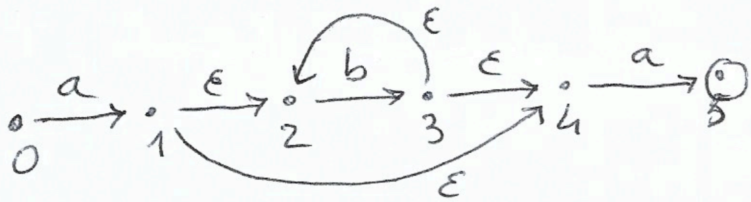
$$\begin{tikzcd}
\boxed{\text{Espr. Regolare}} \arrow[r] & \boxed{\text{NFA}} \arrow[d] \\
\boxed{\text{Gram. Regolare}} \arrow[u] & \boxed{\text{DFA}} \arrow[l]
\end{tikzcd}$$

commuta, ovvero partendo da una espr. regolare s, ritorno su una espr. regolare s' tale che

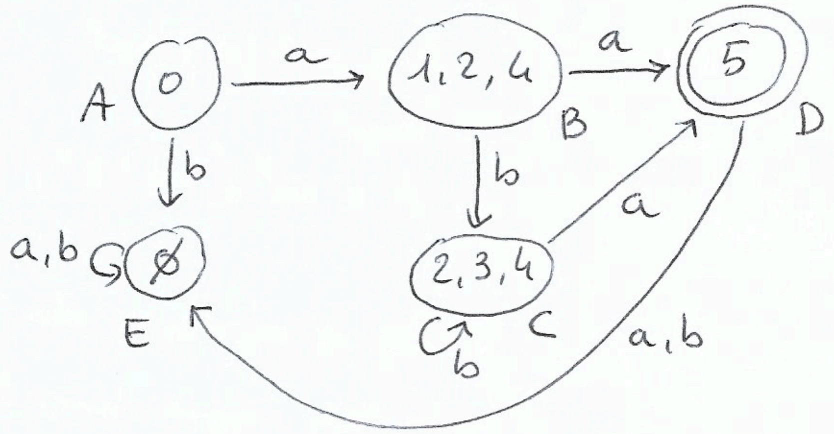
$$s \simeq s'$$

Altro Esempio

1) $a\,b^*a$ Costruire NFA associato



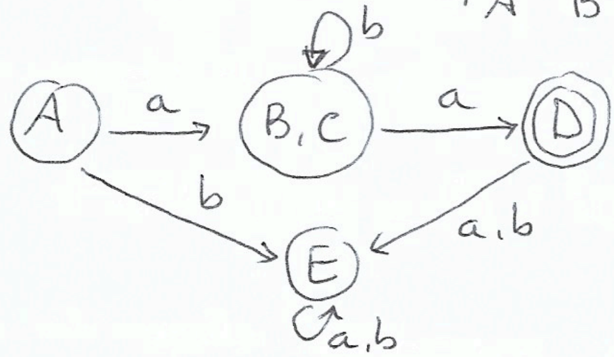
2) Costruire DFA associato



	A	B	C	D
B	X_1			
C	X_1			
D	X_0	X_0	X_0	
E	X_2	X_1	X_1	X_0

$B \sim C$

3) Minimizzare



4) Gr. regolare

$A \rightarrow aB \mid \cancel{bE}$

$B \rightarrow bB \mid \cancel{aD} \mid a$

$\cancel{D} \rightarrow \cancel{aE} \mid \cancel{bE}$

$\cancel{E} \rightarrow \cancel{aE} \mid \cancel{bE}$

5) Semplificare la gram.

$A \rightarrow aB$

$B \rightarrow bB \mid a$

6) Espressione regolare

$B \simeq b^*a$

$A \simeq a\,b^*a$

(per caso uguale a quella di partenza!)

Esercizio: ricavare la gram. regolare con la regola alternativa, poi semplificarla, poi ricavare l'espr. regolare

Automa Minimo

Dato un DFA $M = (\Sigma, Q, \delta, q_0, F)$, l'automa minimo equivalente $M_{min} = (\Sigma, Q_{min}, \delta_{min}, [q_0], F_{min})$ è dato da:

- $Q_{min} = \{ [q] \mid q \in Q \}$ con $[q] = \{ q' \in Q \mid q \sim q' \}$

(gli stati di M_{min} sono classi di equivalenza di stati di M)

- $\delta_{min}([q], a) = [\delta(q, a)]$
- $F_{min} = \{ [q] \mid q \in F \}$

Oss: Non esistono 2 stati distinti in M_{min} che siano tra loro equivalenti: $[q] \neq [q'] \Rightarrow q \not\sim q'$

Notazione (già vista per M, ed estesa a M_{min})

DFA M	DFA M_{min}
$\hat{\delta}: Q \times \Sigma^* \to Q$	$\hat{\delta}_{min}: Q_{min} \times \Sigma^* \to Q_{min}$
$\hat{\delta}(q, \varepsilon) = q$	$\hat{\delta}_{min}([q], \varepsilon) = [q]$
$\hat{\delta}(q, xa) = \delta(\hat{\delta}(q, x), a)$	$\hat{\delta}_{min}([q], xa) = \delta_{min}(\hat{\delta}_{min}([q], x), a)$
$w \in L[M]$ sse $\hat{\delta}(q_0, w) \in F$	$w \in L[M_{min}]$ sse $\hat{\delta}_{min}([q_0], w) \in F_{min}$

Teorema

Dato un DFA $M = (\Sigma, Q, \delta, q_0, F)$, l'automa $M_{min} = (\Sigma, Q_{min}, \delta_{min}, [q_0], F_{min})$ riconosce lo stesso linguaggio di M, ed ha il minimo numero di stati tra tutti gli automi deterministici per questo linguaggio.

Dimostrazione:

(1) M_{min} è ben definito, cioè la definizione di δ_{min} non dipende dallo specifico stato scelto per rappresentare la classe di equivalenza.
Infatti se $q \sim q'$, allora

1) $[q] = [q']$ e

2) $\delta_{min}([q], a) = [\delta(q, a)]$
$= [\delta(q', a)]$
$= \delta_{min}([q'], a)$

se non fosse vero, allora $\delta(q, a)$ e $\delta(q', a)$ sarebbero distinguibili, e quindi pure q e q'

(2) Per dimostrare che $L[M] = L[M_{min}]$, dimostriamo che

$$\hat{\delta}(q_0, w) = r \quad \text{sse} \quad \hat{\delta}_{min}([q_0], w) = [r]$$

ovvero che

$$\hat{\delta}_{min}([q_0], w) = [\hat{\delta}(q_0, w)]$$

Si fa la dimostrazione per induzione su $|w|$

Base: $|w| = 0$ cioè $w = \varepsilon$

$$\hat{\delta}(q_0, \varepsilon) = q_0 \qquad \hat{\delta}_{min}([q_0], \varepsilon) = [q_0] = [\hat{\delta}(q_0, \varepsilon)]$$

ok

Passo Induttivo $w = xa$

$$\hat{\delta}(q_0, xa) = \delta(\hat{\delta}(q_0, x), a)$$

e

$$\hat{\delta}_{min}([q_0], xa) \underset{\substack{\text{per def.} \\ \text{di } \hat{\delta}_{min}}}{=} \delta_{min}(\underbrace{\hat{\delta}_{min}([q_0], x)}_{\text{per ipotesi induttiva}}, a)$$

$$= \delta_{min}(\overbrace{[\hat{\delta}(q_0, x)]}^{=}, a)$$

per def. di δ_{min} $\longrightarrow$ $= [\delta(\hat{\delta}(q_0, x), a)]$

per def. di $\hat{\delta}$ $\longrightarrow$ $= [\hat{\delta}(q_0, xa)]$ ok

Quindi

$w \in L[M]$ sse $\hat{\delta}(q_0, w) = r \in F$

$$\Updownarrow$$

$\hat{\delta}_{min}([q_0], w) = [r] \in F_{min}$ sse $w \in L[M_{min}]$

(3) Rimane da dimostrare che è minimo, ovvero che un qualunque altro automa deterministico N non può avere meno stati

Supponiamo ~~esista~~ esista un DFA N tale che $L[N] = L[M_{min}]$, ma con un numero di stati inferiore a quello di M_{min}.

<u>Oss 1:</u> Gli stati iniziali di N e M_{min} devono essere equivalenti (nell'automa $N \cup M_{min}$) perché $L[N] = L[M_{min}]$

si suppone che N e M_{min} siano disgiunti!

<u>Oss 2:</u> Se p in Q_{min} e q in N sono equivalenti, nell'automa $(N \cup M_{min})$, allora sono equivalenti anche i loro successori $\forall a \in \Sigma$.

<u>Oss 3:</u> - N non ha stati irraggiungibili dal suo iniziale (altrimenti potrei costruire N' con ancora meno stati)

- M_{min} non ha stati irraggiungibili per costruzione

$\Rightarrow$ ogni stato di M_{min} è equivalente ad almeno uno stato di N (per <u>OSS</u> (2))

<u>Oss 4:</u> Poiché N ha meno stati di M_{min}, due stati p e p' di M_{min} devono essere equivalenti ad uno stesso stato q di N. Ma la relazione di eq. $\sim$ è transitiva $\Rightarrow$ p e p' devono essere equivalenti!

Ma questo è impossibile!

Per costruzione di M_{min}, non ci sono 2 stati diversi in M_{min} equivalenti tra loro!