Corso di Laurea in Sicurezza dei sistemi e delle reti informatiche

# Fondamenti di informatica per la sicurezza

anno accademico 2005–2006

docente: Stefano Ferrari

## 20.01.2006 — Soluzione del secondo compitino — vers. C

valutazioni **1** (4) _____ **2** (4) _____ **3** (4) _____ **4** (6) _____ **5** (6) _____ **6** (8) _____

**Cognome** _____

**Nome** _____

**Matricola** _____ **Firma** _____

### Esercizio 1

Siano dati i linguaggi $L_1$ e $L_2$:

- $L_1 = \{z, x, yx\}$
- $L_2 = \{x, y\}$

Descrivere i linguaggi:

a) $L_3 = L_1 \cap L_2$

b) $L_4 = L_1 \cup L_2$

c) $L_5 = L_1 L_2$

d) $L_6 = {L_1}^2$

e) $L_7 = {L_2}^* {L_1}^*$

f) $L_8 = ({L_1}^* L_2)^3$

Per quegli insiemi di cui sia troppo lungo (o impossibile) dare una descrizione estensionale, elencare almeno tre elementi, indicando le caratteristiche degli elementi che li compongono. In particolare, chiarire se la stringa vuota $\epsilon$ appartiene al linguaggio.

### Soluzione

a) $L_3 = L_1 \cap L_2 = \{x\}$
Gli insiemi $L_1$ e $L_2$ hanno un unico elemento in comune, $x$.

b) $L_4 = L_1 \cup L_2 = \{z, x, yx, y\}$
L'elemento comune, $x$, non deve essere ripetuto due volte.

c) $L_5 = L_1 L_2 = \{zx, zy, xx, xy, yxx, yxy\}$

d) $L_6 = {L_1}^2 = \{zz, zx, zyx, xz, xx, xyx, yxz, yxx, yxyx\}$

e) $L_7 = {L_2}^* {L_1}^*$
L'insieme $L_7$ è dato dalle stringhe formate come concatenazione di un numero arbitrario (eventualmente nullo) di elementi di $L_2$ seguito da una concatenazione di un numero arbitrario (eventualmente nullo) di elementi di $L_1$. Poiché sia ${L_1}^*$ che ${L_2}^*$ sono composti da infiniti elementi, anche $L_7$ avrà infiniti elementi. L'insieme $\{\epsilon, xyxyy, xxzzyx, xxxzyx\}$ è un sottoinsieme di $L_7$.

f) $L_8 = ({L_1}^* L_2)^3$
L'insieme $L_8$ è formato dalla concatenazione di 3 stringhe composte dalla numero arbitrario (eventualmente nullo) di elementi di $L_1$ e da un elemento di $L_2$. Pertanto, $L_8$ è composto da infiniti elementi, ma, poiché deve sempre esserci un elemento di $L_2$ ed $L_2$ non contiene la stringa vuota, $\epsilon$ non appartiene a $L_8$. L'insieme $\{yyy, zxzxzxzyxyyxx, zyyxy\}$ è un sottoinsieme di $L_8$.

### Esercizio 2

Sia data la seguente grammatica, $G = \langle T, V, P, S\rangle$, definita su $\Sigma = \{a, b, c, d\}$:

- insieme dei simboli terminali, $T$: $T = \Sigma$
- insieme dei metasimboli, $V$: $V = \{K, H\}$
- insieme delle regole di produzione, $P$: $P = \{S ::= K, K ::= b|aH|cH, H ::= d|cK|bH\}$

Quali fra le seguenti stringhe vengono generate da $G$?

a) *cbca*

b) *acaa*

c) *ccad*

d) *abcad*

e) *accbcba*

Riportare la successione di regole da applicare per la generazione di tali stringhe e le stringhe parziali ottenute, spiegando perché non si possono ottenere le stringhe che eventualmente non risultassero appartenere al linguaggio generato da $G$.

## Soluzione

a)

| *cbca* | |
|---|---|
| | $S$ |
| $S ::= K$ | $K$ |
| $K ::= cH$ | $cH$ |
| $H ::= bH$ | $cbH$ |
| $H ::= cK$ | $cbcK$ |
| $K ::= aH$ | $cbcaH$ |

Non è possibile eliminare il metasimbolo $H$ senza aggiungere un altro simbolo.

La stringa *cbca* non è generata da $G$: $cbca \notin L(G)$.

b)

| *acaa* | |
|---|---|
| | $S$ |
| $S ::= K$ | $K$ |
| $K ::= aH$ | $aH$ |
| $H ::= cK$ | $acK$ |
| $K ::= aH$ | $acaH$ |

Non esiste regola che generi il simbolo $a$ dal metasimbolo $H$.

La stringa *acaa* non è generata da $G$: $acaa \notin L(G)$.

c)

| *ccad* | |
|---|---|
| | $S$ |
| $S ::= K$ | $K$ |
| $K ::= cH$ | $cH$ |
| $H ::= cK$ | $ccK$ |
| $K ::= aH$ | $ccaH$ |
| $H ::= d$ | $ccad$ |

La stringa *ccad* è generata da $G$: $ccad \in L(G)$.

d)

| *abcad* | |
|---|---|
| | $S$ |
| $S ::= K$ | $K$ |
| $K ::= aH$ | $aH$ |
| $H ::= bH$ | $abH$ |
| $H ::= cK$ | $abcK$ |
| $K ::= aH$ | $abcaH$ |
| $H ::= d$ | $abcad$ |

La stringa *abcad* è generata da $G$: $abcad \in L(G)$.

e)

| *accbcba* | |
|---|---|
| | $S$ |
| $S ::= K$ | $K$ |
| $K ::= aH$ | $aH$ |
| $H ::= cK$ | $acK$ |
| $K ::= cH$ | $accH$ |
| $H ::= bH$ | $accbH$ |
| $H ::= cK$ | $accbcK$ |
| $K ::= b$ | $accbcb$ |

La stringa generata non coincide con la stringa data, *accbcba*, e non può essere ulteriormente estesa, per mancanza di metasimboli.

La stringa *accbcba* non è generata da $G$: $accbcba \notin L(G)$.

## Esercizio 3

Sia dato il seguente automa a stati finiti, $A$, $A = \langle Q, \Sigma, \delta, q_0, F \rangle$:

- insieme degli stati, $Q$: $Q = \{q_0, q_1, q_2, q_3\}$
- alfabeto di input, $\Sigma$: $\Sigma = \{a, b, c, d, e\}$
- funzione di transizione $\delta$:

| | $a$ | $b$ | $c$ | $d$ | $e$ |
|---|---|---|---|---|---|
| $q_0$ | $q_2$ | $q_0$ | $q_1$ | $q_0$ | $q_3$ |
| $q_1$ | $q_1$ | $q_2$ | $q_1$ | $q_1$ | $q_1$ |
| $q_2$ | $q_3$ | $q_2$ | $q_0$ | $q_0$ | $q_2$ |
| $q_3$ | $q_2$ | $q_1$ | $q_2$ | $q_3$ | $q_1$ |

- stato iniziale, $q_0$
- insieme di stati finali, $F$: $F = \{q_1\}$

Indicare:

a) quattro stringhe accettate da $A$

b) quattro stringhe rifiutate da $A$

### Soluzione

a) quattro stringhe accettate da $A$:

- *ecabac*
- *ebdce*
- *cede*
- *badc*

b) quattro stringhe rifiutate da $A$:

- *ddea*
- *aed*
- *ddab*
- *aa*

## Esercizio 4

Modellare, tramite un automa a stati finiti deterministico, il funzionamento di una cassetta della posta.

Una cassetta della posta ha capienza tale da poter ospitare al più tre buste o un plico e una busta. La cassetta è dotata di uno sportello che deve essere aperto per poter prelevare il contenuto della cassetta. L'inserimento di plichi e buste può invece essere effettuato anche a sportello chiuso.

Ipotizzare che non si possano verificare contemporaneamente più azioni. Modellare l'automa in modo che esso accetti solo le stringhe che descrivono il normale utilizzo della cassetta della posta. In particolare, individuare possibili situazioni fisicamente irrealizzabili o pericolose e formalizzarle in modo che l'automa rifiuti le successioni di azioni che porterebbero la cassetta in tali situazioni.

Stati e simboli riportati o suggeriti nel testo sono solo indicativi: possono essere modificati, ridotti ed estesi a secondo delle esigenze del progetto.

### Soluzione

L'automa deve modellare un sistema fisico. L'insieme dei simboli di input modella quindi gli stimoli che il sistema riceve dall'esterno (o le azioni che esso subisce) e gli stati descrivono le situazioni in cui il sistema viene a trovarsi.

Questo permette di vedere l'automa come un simulatore del sistema in esame: l'automa deve accettare le stringhe che rappresentano le sequenze di stimoli (o azioni) fisicamente realizzabili oppure quelle che rappresentano una sequenza di eventi di particolare interesse.

Il sistema del problema è descrivibile come insieme di due sottosistemi indipendenti: lo sportello e l'interno. Lo sportello può trovarsi in due stati (aperto o chiuso), mentre l'interno della cassetta può trovarsi in 5 stati (vuota, con una busta, con due buste, con tre buste, e con una busta e un plico).

Gli stati del carillon si possono derivare dalle combinazioni degli stati dei due sottosistemi. Una prima semplificazione può essere operata considerando che dalle specifiche si può dedurre che i plichi occupino il volume di due buste e che ogni busta occupi un terzo del volume della casetta. Quindi è possibile ridurre a 4 gli stati dell'interno della cassetta, considerando il numero di terzi di cassetta occupati (0, 1, 2, e 3) come stati dell'interno.

Pertanto, l'insieme degli stati, $Q$, può essere:

$$Q = \{a0,\ a1,\ a2,\ a3,\ c0,\ c1,\ c2,\ c3\}$$

dove la lettera $a$ indica se lo sportello è aperto, in contrapposizione a $c$ che indica che lo sportello è chiuso, mentre i numeri 0–3 indicano il numero di terzi di cassetta disponibili.

Le azioni che possono essere effettuate sul sistema sono l'apertura e la chiusura dello sportello, il prelevamento della posta, l'introduzione di una busta e l'introduzione di un plico.

Pertanto, insieme dei simboli, $\Sigma$, può essere:

$$\Sigma = \{a,\ c,\ s,\ b,\ p\}$$

dove $a$ e $c$ indicano, rispettivamente, l'azione di apertura e di chiusura dello sportello, $s$ indica lo svuotamento della cassetta, $b$ indica l'introduzione di una busta e $p$ quella di un plico.

Le specifiche descrivono i seguenti comportamenti:

- il prelevamento della posta può avvenire solo a sportello aperto;
- il tentativo di introdurre posta a cassetta piena va ritenuto un caso di malfunzionamento.

Le specifiche non impongono particolari condizioni per l'introduzione della corrispondenza. Pertanto, si assumerà che bustte e plichi possano essere introdotti in ogni momento, indipendentemente dallo stato dello sportello.

Poiché esiste almeno una situazione ritenuta non accettabile, è opportuno aggiungere un altro stato, *errore*, tale per cui una volta raggiunto

non lo si possa più lasciare. Questa caratteristica formalizza il fatto che la situazione di errore è irreversibile, cioè non esiste una sequenza di azioni che permetta di riassorbire una situazione non accettabile.

Si può inoltre ipotizzare che il tentativo di apertura (chiusura) dello sportello già aperto (chiuso) generi errore. In tal caso, l'automa viene portato nello stato *errore*.

Ogni sequenza di azioni che non comporti il raggiungimento dello stato *errore* rappresenta il normale funzionamento della cassetta della posta. Pertanto, qualsiasi sequenza di simboli che non porti nello stato *errore* deve venire accettata, e, quindi, tutti gli stati tranne *errore* compongono l'insieme degli stati finali, $F$.

Si può ipotizzare che lo stato iniziale sia quello relativo alla cassetta vuota e con sportello chiuso, $c0$.

Quindi, l'insieme degli stati, $Q$, qui utilizzato sarà:

$$Q = \{a0,\ a1,\ a2,\ a3,\ c0,\ c1,\ c2,\ c3,\ errore\}$$

Con le ipotesi fatte, dovrebbero essere accettate, per esempio, le seguenti sequenze di azioni: *bbbasc*, *apbc*. Al contrario, verrebbero rifiutate, tra le altre, le seguenti sequenze di azioni: *pbb*, *ps*, *abbcs*. Va notato che aggiungendo un qualsiasi suffisso ad una stringa rifiutata, si ottiene sempre una stringa rifiutata: se una certa sequenza di azioni porta in uno stato non accettabile, qualsiasi sequenza di azioni ad essa successiva non può renderla accettabile.

La tabella delle transizioni, $\delta : Q \times \Sigma \to Q$ può essere quella riportata in Tabella 1.

Modellazioni differenti possono essere ottenute con differenti scelte progettuali. Ad esempio, può essere considerata la rimozione della singola busta e plico: in questo caso il numero di stati e di simboli cresce, soprattutto se si considera la corrispondenza accumulata come facente parte di una pila (si rimuove in ordine inverso a quello di introduzione).

## Esercizio 5

Sia data l'espressione regolare $E$, definita su $\Sigma = \{a,\ b,\ c\}$:

- $E = (a^3c^*b)^* + (c^2b^*a)^*$

Quali fra le seguenti stringhe vengono descritte da $E$?

a) *aaacbaaacccbaaab*

b) *ccaccbbbaa*

c) *ccbbaccaccba*

d) *aaabaaaccbaaab*

e) *aaabaaabc*

f) *ccbbbaccba*

### Soluzione

Le espressioni regolari denotano degli insiemi di stringhe. In tal senso, possiamo applicare l'operatore di relazione insiemistica $\subseteq$ alle espressioni regolari per indicare che l'insieme denotato da un'espressione contiene l'insieme denotato da una seconda espressione regolare. Per esempio, $E_1 \subseteq E_2$ significa che tutte le stringhe descritte da $E_1$ sono descritte anche da $E_2$.

Ricordando che l'espressione regolare $s$ descrive l'insieme di stringhe composto dalla sola $s$, $\{s\}$, si può dimostrare che tale stringa viene descritta da un'espressione regolare $E$ derivando una catena di inclusioni del tipo $s \subseteq E_1 \subseteq E_2 \subseteq \cdots \subseteq E_k \equiv E$.

Osserviamo innanzitutto l'espressione regolare $E$ è costituita dall'unione di due sottoespressioni: $E_1 = (a^3c^*b)^*$ e $E_2 = (c^2b^*a)^*$. Quindi, le stringhe descritte da $E$ dovranno essere descritte da $E_1$ o, in alternativa, da $E_2$. Questa premessa semplificherà la spiegazione delle soluzioni di seguito riportate.

a) *aaacbaaacccbaaab*
$aaacbaaacccbaaab = (aaacb)(aaacccb)(aaab) = (a^3cb)(a^3c^3b)(a^3b) \subseteq (a^3c^*b)(a^3c^*b)(a^3c^*b) \subseteq (a^3c^*b)^3 \subseteq (a^3c^*b)^* \subseteq (a^3c^*b)^* + (c^2b^*a)^*$

La stringa *aaacbaaacccbaaab* viene descritta da $E$: $aaacbaaacccbaaab \in \mathcal{L}(E)$.

b) *ccaccbbbaa*

Iniziando per $c$, la stringa data non può essere descritta da $E_1$ e, terminando per $aa$, non può nemmeno essere descritto da $E_2$.

Pertanto, la stringa *ccaccbbbaa* non viene descritta da $E$: $ccaccbbbaa \notin \mathcal{L}(E)$.

c) *ccbbaccaccba*
$ccbbaccaccba = (ccbba)(cca)(ccba) = (c^2b^2a)(c^2a)(c^2ba) \subseteq (c^2b^*a)(c^2b^*a)(c^2b^*a) \subseteq (c^2b^*a)^3 \subseteq (c^2b^*a)^* \subseteq (a^3c^*b)^* + (c^2b^*a)^*$

| $\delta$ | $a$ | $c$ | $s$ | $b$ | $p$ |
|---|---|---|---|---|---|
| $a0$ | $errore$ | $c0$ | $a0$ | $a1$ | $a2$ |
| $a1$ | $errore$ | $c1$ | $a0$ | $a2$ | $a3$ |
| $a2$ | $errore$ | $c2$ | $a0$ | $a3$ | $errore$ |
| $a3$ | $errore$ | $c3$ | $a0$ | $errore$ | $errore$ |
| $c0$ | $a0$ | $errore$ | $errore$ | $c1$ | $c2$ |
| $c1$ | $a1$ | $errore$ | $errore$ | $c2$ | $c3$ |
| $c2$ | $a2$ | $errore$ | $errore$ | $c3$ | $errore$ |
| $c3$ | $a3$ | $errore$ | $errore$ | $errore$ | $errore$ |
| $errore$ | $errore$ | $errore$ | $errore$ | $errore$ | $errore$ |

Tabella 1: Tabella delle transizioni dell'automa dell'esercizio 4.

La stringa *ccbbaccaccba* viene descritta da $E$: $ccbbaccaccba \in \mathcal{L}(E)$.

d) *aaabaaaccbaaab*
$aaabaaaccbaaab = (aaab)(aaaccb)(aaab) = (a^3b)(a^3c^2b)(a^3b) \subseteq (a^3c^*b)(a^3c^*b)(a^3c^*b) \subseteq (a^3c^*b)^3 \subseteq (a^3c^*b)^* \subseteq (a^3c^*b)^* + (c^2b^*a)^*$

La stringa *aaabaaaccbaaab* viene descritta da $E$: $aaabaaaccbaaab \in \mathcal{L}(E)$.

e) *aaabaaabc*

Iniziando per $a$, la stringa data non può essere descritta da $E_2$, e, terminando per $c$, non può essere descritta da $E_1$.

Pertanto, la stringa *aaabaaabc* non viene descritta da $E$: $aaabaaabc \notin \mathcal{L}(E)$.

f) *ccbbbaccba*
$ccbbbaccba = (ccbbba)(ccba) = (c^2b^3a)(c^2ba) \subseteq (c^2b^*a)(c^2b^*a) \subseteq (c^2b^*a)^2 \subseteq (c^2b^*a)^* \subseteq (a^3c^*b)^* + (c^2b^*a)^*$

La stringa *ccbbbaccba* viene descritta da $E$: $ccbbbaccba \in \mathcal{L}(E)$.

## Esercizio 6

Indicare una espressione regolare (non banale) definita su $\Sigma = \{a, b, c\}$ che descriva le seguenti stringhe:

- *acbbccacca*
- *acacbbacca*
- *acbacba*
- *bbb*

ma non le seguenti:

- *caccbcca*
- *baabaa*
- *bbbca*
- *acbca*

## Soluzione

Le stringhe da accettare hanno nella parte iniziale le stringhe $b$ e $ac$. In particolare, almeno tre occorrenze di queste stringhe costituiscono il suffisso delle stringhe da accettare. Infatti:

- *acbb*ccacca: inizia con $ac$ seguito da due occorrenze di $b$;
- *acacbb*acca: inizia con due occorrenze di $ac$ seguite da $b$;
- *acbac*ba inizia con due occorrenze di $ac$ intercalate con $b$;
- *bbb*: inizia con (anzi, è formato da) due occorrenze di $ac$.

Questa caratteristica può essere descritta dall'espressione regolare $(b + ac)^3(a + b + c)^*$.

Delle stringhe del secondo gruppo solo una, *bbbca*, viene descritta da tale espressione regolare; riguardo le altre, infatti, si può notare che:

- *caccbcca*: inizia per $c$;
- *baabaa*: dopo una $b$ iniziale, ha una doppia $a$;
- *acbca*: inizia con $ac$ seguito da $b$, ma poi prosegue con $ca$.

Per escludere la stringa rimanente, bisogna rendere un po' più specifica l'espressione regolare fin qui individuata. Ciò può essere fatto osservando che nella parte finale delle stringhe da accettare, il simbolo $c$, se compare, è sempre a coppie. In altri termini, il simbolo $c$ compare

solo all'interno di sottostringhe $cc$. Pertanto, l'espressione $(b{+}ac)^3(a{+}b{+}cc)^*$ risolve il problema dato.

Altre espressioni regolari che rispettano le specifiche del problema sono:

- $(b + ac)^3(c^2a + ba)^*$
- $b^* + ((a + aca)((c^*b^*)^2a)^2)$
- $(a + b)(a + b + c)^*(b + cca + cba)$
- $(acb^* + c^2a + a)^* + b^3$
- $(ac^*b^* + c^2a + b^3)^*$
- $(ac)^*b(ac)^*b(a + c)^*b^*$