Università degli Studi di Milano

Corso di Laurea in Sicurezza dei sistemi e delle reti informatiche

Fondamenti di informatica per la sicurezza

anno accademico 2005–2006 docente: Stefano Ferrari

20.04.2006 — Soluzione della seconda parte — versione A

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:

- $L_1 = \{a, b, ba\}$
- $L_2 = \{ab, c\}$

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) $L_6 = {L_1}^3$

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

f) $L_8 = (L_2 L_1)^*$

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 ϵ appartiene al linguaggio.

Soluzione

a) $L_3 = L_1 \cap L_2 = \emptyset$
Gli insiemi L_1 e L_2 non hanno elementi in comune, quindi la loro intersezione è vuota.

Nota: L'insieme vuoto $\emptyset$ è diverso dall'insieme costituito dalla sola stringa vuota, $\{\epsilon\}$.

b) $L_4 = L_1 \cup L_2 = \{a, ab, b, ba, c\}$

c) $L_5 = L_1 L_2 = \{aab, ac, baab, bab, bac, bc\}$

d) Il linguaggio $L_6 = L_1^3$ ha 26 elementi. L'insieme $\{aaa, aab, bbab, bbaba, bbb, bbba\}$ è un sottoinsieme di L_6.
Gli elementi che possono essere ottenuti in più di un modo devono essere riportati solo una volta.

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

f) $L_8 = (L_2 L_1)^*$
L'insieme L_8 è formato dalla concatenazione di un numero arbitrario (eventualmente nullo) di stringhe composte da un elemento di L_2 e da un elemento di L_1. Pertanto, L_8 è composto da infiniti elementi. L'insieme $\{\epsilon, cb, cacbaabaabb\}$ 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: $P = \{S ::= H, K ::= a|Hc|Hb, H ::= b|Ka|Hd\}$

Quali fra le seguenti stringhe vengono generate da G?

a) $bcad$

b) $baadd$

c) $cdca$

d) $adba$

e) $aacaba$

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)

$bcad$	
	S
$S ::= H$	H
$H ::= Hd$	Hd
$H ::= Ka$	Kad
$K ::= Hc$	$Hcad$
$H ::= b$	$bcad$

La stringa $bcad$ è generata da G: $bcad \in \mathcal{L}(G)$.

b)

$baadd$	
	S
$S ::= H$	H
$H ::= Hd$	Hd
$H ::= Hd$	Hdd
$H ::= Ka$	$Kadd$
$K ::= a$	$aadd$

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

La stringa $baadd$ non è generata da G: $baadd \notin \mathcal{L}(G)$.

c)

$cdca$	
	S
$S ::= H$	H
$H ::= Ka$	Ka
$K ::= Hc$	Hca
$H ::= Hd$	$Hdca$

Non esiste regola che permetta di ottenere il simbolo c dal metasimbolo H.

La stringa $cdca$ non è generata da G: $cdca \notin \mathcal{L}(G)$.

d)

$adba$	
	S
$S ::= H$	H
$H ::= Ka$	Ka
$K ::= Hb$	Hba
$H ::= Hd$	$Hdba$
$H ::= Ka$	$Kadba$

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

La stringa $adba$ non è generata da G: $adba \notin \mathcal{L}(G)$.

e)

$aacaba$	
	S
$S ::= H$	H
$H ::= Ka$	Ka
$K ::= Hb$	Hba
$H ::= Ka$	$Kaba$
$K ::= Hc$	$Hcaba$
$H ::= Ka$	$Kacaba$
$K ::= a$	$aacaba$

La stringa $aacaba$ è generata da G: $aacaba \in \mathcal{L}(G)$.

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, Σ: $\Sigma = \{a, b, c, d, e\}$
- funzione di transizione δ:

	a	b	c	d	e
q_0	q_1	q_1	q_1	q_3	q_1
q_1	q_2	q_3	q_1	q_0	q_0
q_2	q_3	q_0	q_1	q_2	q_3
q_3	q_3	q_2	q_0	q_0	q_2

- 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:

- $baace$
- $daaadb$
- $accc$

- *cbdcc*

b) quattro stringhe rifiutate da A:

- *daaa*
- *abcdc*
- *dbeac*
- *bcaad*

Esercizio 4

Modellare, tramite un automa a stati finiti deterministico, il funzionamento di una lavatrice.

Una lavatrice è dotata di uno sportello che permette l'accesso al cestello e di un pulsante di attivazione. Nel normale funzionamento, l'utente apre lo sportello, inserisce i panni da lavare, chiude lo sportello e attiva la lavatrice. Al termine del ciclo di lavaggio, l'utente svuota il cestello.

Ipotizzando che l'utente sia interessato solo a lavare camicie e pantaloni, considerare un paio di pantaloni equivalente per peso ed ingombro a due camicie.

Il cestello della lavatrice ha una capienza massima di cinque camicie, ma in fase di lavaggio, puo' sopportare il peso solo di quattro camicie.

Attivare la lavatrice senza aver chiuso lo sportello non ha effetti, così come cercare di aprire lo sportello con la lavatrice in funzione (un dispositivo di blocco impedisce tale operazione).

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

La macchina che deve essere descritta dall'automa è composta da alcuni sottosistemi, responsabili di particolari funzioni: il cestello, lo sportello e il pulsante di attivazione. L'insieme di stati in cui la lavatrice può trovarsi sono descritti dalle combinazioni possibili dei suoi sottosistemi.

Lo stato in cui si trova il cestello può essere descritto dal volume di carico e dal tipo di indumenti caricati. Per semplificare la trattazione, si può considerare solo il volume caricato, tenendo presente che un paio di pantaloni occupa lo stesso volume di due camicie. La sua capienza è pari a 5 camicie, quindi il cestello può trovarsi in 6 stati: vuoto, una camicia, due camicie, e così via fino a cinque camicie. Lo sportello può essere aperto o chiuso, quindi due stati sono sufficienti per descrivere tale sottosistema. Infine, la lavatrice può essere attiva o a riposo: anche questa situazione può essere descritta da due stati.

Quindi lo stato della lavatrice potrebbe essere descritto da $6 \times 2 \times 2 = 24$ stati. Tuttavia, le specifiche consentono di ridurre il numero di stati necessari. La lavatrice non può essere attiva e con lo sportello aperto. Quindi, le combinazioni degli stati dei sottosistemi sportello e pulsante di attivazione possono essere ridotti a 3: sportello aperto, sportello chiuso (ma lavatrice a riposo) e lavatrice attiva (con sportello chiuso).

Inoltre le specifiche dicono che se la lavatrice non può sopportare un lavaggio con un carico equivalente a 5 camicie, quindi deve essere previsto anche uno stato di errore.

Quindi, l'insieme degli stati, Q, può quindi essere:

$$\begin{aligned} Q \;=\; & \{a_0,\, a_1,\, a_2,\, a_3,\, a_4,\, a_5,\, c_0,\, c_1,\, c_2,\, c_3,\, c_4, \\ & c_5,\, l_0,\, l_1,\, l_2,\, l_3,\, l_4,\, \mathit{errore}\} \end{aligned}$$

dove la prima lettera indica lo stato relativo allo sportello ed al pulsante di attivazione (a per sportello aperto, c per sportello aperto, e l per lavaggio), mentre il numero in pedice indica lo stato del cestello, espresso in camicie caricate.

Lo stato *errore* è 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.

Le azioni che possono essere effettuate sono:

- l'inserimento di un capo (camicia o pantalone);
- l'estrazione di un capo (camicia o pantalone);
- l'apertura dello sportello;
- la chiusura dello sportello;
- la attivazione della lavatrice;
- lo spegnimento della lavatrice.

Per via della semplificazione scelta di non tener conto del tipo di indumenti caricati, ma solo del loro volume, non è possibile distinguere tra un carico di due camicie e di un paio di pantaloni. Pertanto, anche l'azione di scarico deve tenere conto di tale semplificazione: vi saranno due azioni di scarico, s_1 e s_2, le quali avranno l'effetto di abbassare il carico del cestello di una quantità pari a, rispettivamente, una e due camicie.

Le specifiche non chiariscono cosa succeda se si cerca di prelevare una quantità di carico non (totalmente) presente nel cestello. Si può ipotizzare che questa azione non abbia conseguenze. Allo stesso modo, si può ipotizzare che il tentativo di aprire lo sportello quando esso sia già aperto, o di chiuderlo quando esso sia già chiuso, non abbia alcun effetto.

L'insieme dei simboli, Σ, può essere:

$$\Sigma = \{i_c,\, i_p,\, s_1,\, s_2,\, a,\, c,\, o,\, t\}$$

dove i_c e i_p rappresentano rispettivamente l'inserimento di una camicia e di un paio di pantaloni, s_1 e s_2, l' estrazione dal cestello di una quantità pari a, rispettivamente, una e due camicie, a e c rappresentano rispettivamente l'apertura e la chiusura dello sportello, e, infine, o e t rappresentano rispettivamente l'attivazione e lo spegnimento della lavatrice.

L'insieme dei simboli può essere esteso aggiungendo un simbolo che rappresenti l'evento di terminazione della operazione di lavaggio. Tuttavia, tale evento può in prima battuta essere rimpiazzato dall'operazione di spegnimento manuale (t).

Si può inoltre ipotizzare che il tentativo di caricare o scaricare il cestello a sportello chiuso causi un errore irreversibile e quindi comporti la transizione dell'automa nello stato *errore*.

Ogni sequenza di azioni che non comporti il raggiungimento dello stato *errore* rappresenta il normale funzionamento della lavatrice. 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 lavatrice scarica, spenta e con lo sportello chiuso, c_0.

Con le ipotesi fatte, dovrebbero essere accettate, per esempio, le seguenti sequenze di azioni: ai_ccot, $ai_ci_pcas_1co$, $ototai_c$. Al contrario, verrebbero rifiutate, tra le altre, le seguenti sequenze di azioni: i_c, ai_cci_p, $ai_ci_pi_pi_c$. 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.

È possibile rendere l'automa più complesso, modellando, per esempio anche il contenuto del cestello, oltre che il volume dei panni inseriti. In tal modo è possibile modellare in modo più fine anche le operazioni di inserimento ed estrazione dei capi. Tuttavia, questa modellazione richiede un numero maggiore di stati: il numero di stati del cestello cresce da 6 a 11, portando così da 18 a 33 il numero di stati dell'automa.

Un'altra variante possibile riguarda l'insieme degli stati finali, F. Se F fosse ristretto allo stato iniziale (cioè lavatrive scarica, disattiva e con lo sportello chiuso) si modellerebbero le sequenze di azioni tali da riportare la lavatrice nella situazione

Esercizio 5

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

- $E = (a + bc)^2(ba^* + c)^*$

Quali fra le seguenti stringhe vengono descritte da E?

a) *bacbac*

b) *abcccbc*

c) *abcabbb*

d) *aa*

e) *bcbcbccb*

f) *bcabaaac*

δ	i_c	i_p	s_1	s_2	a	c	o	t
a_0	a_1	a_2	a_0	a_0	a_0	c_0	a_0	a_0
a_1	a_2	a_3	a_0	a_0	a_1	c_1	a_1	a_1
a_2	a_3	a_4	a_1	a_0	a_2	c_2	a_2	a_2
a_3	a_4	a_5	a_2	a_1	a_3	c_3	a_3	a_3
a_4	a_5	*errore*	a_3	a_2	a_4	c_4	a_4	a_4
a_5	*errore*	*errore*	a_4	a_3	a_5	c_5	a_5	a_5
c_0	*errore*	*errore*	*errore*	*errore*	a_0	c_0	l_0	c_0
c_1	*errore*	*errore*	*errore*	*errore*	a_1	c_1	l_1	c_1
c_2	*errore*	*errore*	*errore*	*errore*	a_2	c_2	l_2	c_2
c_3	*errore*	*errore*	*errore*	*errore*	a_3	c_3	l_3	c_3
c_4	*errore*	*errore*	*errore*	*errore*	a_4	c_4	l_4	c_4
c_5	*errore*	*errore*	*errore*	*errore*	a_5	c_5	*errore*	c_5
l_0	*errore*	*errore*	*errore*	*errore*	l_0	l_0	l_0	c_0
l_1	*errore*	*errore*	*errore*	*errore*	l_1	l_1	l_1	c_1
l_2	*errore*	*errore*	*errore*	*errore*	l_2	l_2	l_2	c_2
l_3	*errore*	*errore*	*errore*	*errore*	l_3	l_3	l_3	c_3
l_4	*errore*	*errore*	*errore*	*errore*	l_4	l_4	l_4	c_4
errore	*errore*	*errore*	*errore*	*errore*	*errore*	*errore*	*errore*	*errore*

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

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 la concatenazione di due sottoespressioni: $E_1 = (a+bc)^2$ e $E_2 = (ba^*+c)^*$. Quindi, le stringhe descritte da E dovranno obbligatoriamente avere un suffisso descritto da E_2 eventualmente preceduto da un prefisso descritto da E_1 (poiché E_1 descrive anche la stringa vuota). Questa premessa semplificherà la spiegazione delle soluzioni di seguito riportate.

a) *bacbac*

La stringa inizia per ba, ma E_1 non può descrivere stringhe in cui il simbolo a segua il simbolo b.

La stringa *bacbac* non viene descritta da E: $bacbac \notin \mathcal{L}(E)$.

b) *abcccbc*
$abcccbc = (a)(bc)(c)(c)(b)(c) \subseteq (a + bc)^2(c)(c)(ba^*)(c) \subseteq (a + bc)^2(c + ba^*)^4 \subseteq (a + bc)^2(ba^* + c)^*$

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

c) *abcabbb*

La sottoespressione E_1 descrive la stringa *abc* che fa da prefisso per la stringa data, ma la rimanente sottostringa, *abbb* non può essere descritta da E_2 perché contiene un simbolo a non preceduto da un simbolo b.

La stringa *abcabbb* non viene descritta da E: $abcabbb \notin \mathcal{L}(E)$.

d) *aa*
$aa = (a)(a) \subseteq (a + bc)(a + bc) \subseteq (a + bc)^2(ba^* + c)^*$

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

e) *bcbcbccb*
$bcbcbccb = (bc)(bc)(b)(c)(c)(b) \subseteq (a+bc)(a+bc)(ba^*)(c)(c)(ba^*) \subseteq (a + bc)^2(ba^* + c)^4 \subseteq (a + bc)^2(ba^* + c)^*$

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

f) *bcabaaac*
$bcabaaac = (bc)(a)(baaa)(c) \subseteq (a + bc)^2(ba^3)(c) \subseteq (a + bc)^2(ba^*)(c) \subseteq (a + bc)^2(ba^* + c)^2 \subseteq (a + bc)^2(ba^* + c)^*$

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

Esercizio 6

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

- *aabbcbc*
- *abbbabb*
- *aabbbcb*
- *ababccc*

ma non le seguenti:

- *bbabbca*
- *abbaba*
- *abbaabbc*
- *acaacb*

Soluzione

Si può notare che tutte le stringhe da includere contengono solo due simboli a, eventualmente separati da simboli b. I rimanenti simboli sono solo b o c. Queste caratteristica può essere descritta dall'espressione regolare $ab^*a(b + c)^*$. Nessuna delle stringhe del secondo gruppo viene descritta da tale espressione regolare:

- *bbabbca*: non inizia per a;
- *abbaba*: ha tre simboli a;
- *abbaabbc*: ha tre simboli a;
- *acaacb*: ha tre simboli a (e le prime due a sono intervallate da c).

Altre espressioni regolari che rispettano le specifiche:

- $ab^*a(bc^*)^*$;
- $(ab^*)^2(b^*c^*)^*$;
- $a(a + b)^2b(a + b + c)^3$;
- $(a^2 + ab)(a + b + c)^4(b + c)$.

Una descrizione alternativa può essere ricavata notando che tutte le stringhe da includere iniziano con a e sono lunghe 7 simboli, mentre le stringhe da escludere che iniziano per a hanno lunghezza 6 o 8. Pertanto, l'espressione $a(a+b+c)^6$ soddisfa i requisiti.