Università degli Studi di Milano

Corso di Laurea in Sicurezza dei sistemi e delle reti informatiche

Fondamenti di informatica per la sicurezza

anno accademico 2006–2007 docente: Stefano Ferrari

12.01.2007 — Soluzione del secondo compitino — versione D

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:

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

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) $L_8 = ({L_2}^2 L_1)^*$

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) $L_3 = L_1 \cap L_2 = \{b\}$

b) $L_4 = L_1 \cup L_2 = \{a,\ b,\ c,\ cb\}$
Gli elementi comuni non devono essere ripetuti.

c) $L_5 = L_1 L_2 = \{ba,\ bb,\ ca,\ cb,\ cba,\ cbb\}$

d) $L_6 = L_1^2 = \{bb,\ bc,\ bcb,\ cb,\ cbb,\ cbc,\ cbcb,\ cc,\ ccb\}$

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 $\{\epsilon,\ cbbbb,\ babb,\ cca\}$ è un sottoinsieme di L_7.

f) $L_8 = ({L_2}^2 L_1)^*$
L'insieme L_8 è formato dalla concatenazione di un numero arbitrario (eventualmente nullo) di stringhe composte da due elementi di L_2 e da un elemento di L_1. Pertanto, L_8 è composto da infiniti elementi. L'insieme $\{\epsilon,\ abc,\ aacabc\}$ 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: $P = \{S ::= K,\ K ::= c|bH|dK,\ H ::= a|dK|cH\}$

Quali fra le seguenti stringhe vengono generate da G?

a) *bddca*

b) *bccddc*

c) *dbca*

d) *dbddb*

e) *ddbb*

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) *bddca*

	S
$S ::= K$	K
$K ::= bH$	bH
$H ::= dK$	bdK
$K ::= dK$	$bddK$
$K ::= c$	$bddc$

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

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

b) *bccddc*

	S
$S ::= K$	K
$K ::= bH$	bH
$H ::= cH$	bcH
$H ::= cH$	$bccH$
$H ::= dK$	$bccdK$
$K ::= dK$	$bccddK$
$K ::= c$	$bccddc$

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

c) *dbca*

	S
$S ::= K$	K
$K ::= dK$	dK
$K ::= bH$	dbH
$H ::= cH$	$dbcH$
$H ::= a$	$dbca$

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

d) *dbddb*

	S
$S ::= K$	K
$K ::= dK$	dK
$K ::= bH$	dbH
$H ::= dK$	$dbdK$
$K ::= dK$	$dbddK$
$K ::= bH$	$dbddbH$

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

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

e) *ddbb*

	S

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

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_2	q_1	q_2	q_3	q_1
q_1	q_3	q_2	q_0	q_0	q_2
q_2	q_2	q_0	q_1	q_0	q_3
q_3	q_1	q_2	q_1	q_1	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:

- *abcc*
- *eeea*
- *cdde*
- *bcbcb*

b) quattro stringhe rifiutate da A:

- *badb*
- *dabe*
- *acdc*
- *eda*

Esercizio 4

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

Una stazione ferroviaria terminale è collegata mediante un binario alla rete ferroviaria. La stazione è dotata di due binari, ciascuno dei quali può ospitare solo un treno. Uno scambio posto all'ingresso della stazione permette ad un treno entrante di essere diretto su uno dei due binari della stazione.

Ipotizzare che i treni si muovano tutti alla stessa velocità. In particolare, il modello deve tener conto del fatto che un treno non può uscire se un altro treno sta entrando, e viceversa.

Ipotizzare che non si possano verificare contemporaneamente più azioni. Modellare l'automa in modo che esso accetti solo le stringhe che descrivono il normale funzionamento della stazione. In particolare, individuare possibili situazioni fisicamente irrealizzabili o pericolose e formalizzarle in modo che l'automa rifiuti le successioni di azioni che porterebbero la stazione 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 che deve essere descritto dall'automa è composta da alcuni sottosistemi: i due binari, la linea di ingresso alla stazione e lo scambio che collega la linea di ingresso con i binari. L'insieme di stati in cui la stazione può trovarsi sono descritti dalle combinazioni possibili dei suoi sottosistemi (o da un suo sottoinsieme).

Ipotizzando che lo scambio possa assumere solo le due posizioni adatte a deviare i treni nei due binari della stazione, il sottosistema scambio può essere descritto da due stati. I binari della stazione possono ospitare al più un treno ciascuno, quindi anche ciasscuno di essi può essere descritto tramite due stati (binario vuoto e treno presente). La linea di ingresso, invece deve essere descritta da tre stati, per indicare se la linea è libera, se vi è un treno uscente o un treno entrante. Quindi lo stato della stazione potrebbe essere descritto da $2 \times 2 \times 2 \times 3 = 24$ stati.

Poiché nelle specifiche viene descritto almeno un caso di grave malfunzionamento, si deve predisporre anche uno stato in grado di catturare questa situazione: lo stato *errore*.

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

$$\begin{aligned} Q = \{ & ea00,\ ea01,\ ea10,\ ea11,\ eb00,\ eb01,\\ & eb10,\ eb11,\ ua00,\ ua01,\ ua10,\ ua11,\ ub00,\\ & ub01,\ ub10,\ ub11,\ la00,\ la01,\ la10,\ la11,\\ & lb00,\ lb01,\ lb10,\ lb11,\ \mathit{errore}\} \end{aligned}$$

dove la prima lettera indica lo stato della linea di ingresso (treno entrante, e, treno uscente, u, e linea libera, l), la seconda lettera indica lo stato dello scambio (a e b indicano su quale binario lo scambio è impostato) e gli ultimi due numeri indicano il numero di treni presenti sui binari (binario a e binario b, nell'ordine).

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.

Ipotizzando che i treni gestiscano autonomamente le operazioni di fermata, le azioni che possono essere effettuate sono:

- ingresso di un treno;
- modifica della posizione dello scambio;
- uscita del treno dal binario a;
- uscita del treno dal binario b.

Inoltre, le specifiche dicono di tener conto della dinamicità del sistema: se un treno sta entrando, non si può farne uscire un altro, ma bisogna attendere che il treno entrante abbia liberato la linea di ingresso. Per rendere questa caratteristica, si può ipotizzare che le azioni di ingresso ed uscita abbiano una durata temporale tale da far transitare i treni sui vari tratti delle linee considerate. In dettaglio, l'azione di ingresso di un treno farà sì che un treno arrivi dall'esterno e occupi il tratto di linea prima dello scambio. Analogamente, le operazioni di uscita avranno una durata tale per cui il treno utilizzi e disimpegni lo scambio; nello stesso tempo, un eventuale treno che si fosse trovato ad occupare la linea esterna in direzione uscente avrà avuto modo di allontanarsi, liberandola per il treno uscente dal binario.

Va inoltre considerato che possono esserci anche intervalli di tempo durante i quali non avviene nessuna delle azioni sopra descritte. Tuttavia, durante tali periodi, i treni dovranno continuare la loro corsa. Per esempio, se ci fosse solo un treno fermo su binario a, l'azione di uscita dal binario a descriverebbe il suo spostamento dal binario a al tratto di linea esterna. Tuttavia, nell'intervallo di tempo seguente tale treno dovrebbe continuare la sua corsa, liberando anche la linea esterna. Questo comportamento può essere descritto da un'azione particolare che tiene conto dello scorrere del tempo. In alternativa, si dovrebbero considerare una serie di azioni per descrivere nel dettaglio gli spostamenti dei treni sui vari binari.

L'unica azione che può essere considerata istantanea è l'azionamento dello scambio.

Le specifiche non chiariscono cosa succeda se, con lo scambio orientato, per esmpio, sul binario b, esca il treno dal binario a. I possibili esiti sono tre:

- lo scambio è sufficientemente elastico per far passare il treno senza risentirne;
- lo scambio è troppo rigido e si rompe;
- lo scambio è rigido, ma anche resistente, e il treno deraglia.

Il terzo scenario è molto improbabile. Tipicamente, i treni che escono dalla stazione vanno a velocità ridotta, e il primo scenario è il più ragionevole. Tuttavia, in mancanza di specifiche, tutti e tre gli scenari sono ugualmente considerabili.

Un'altra situazione da considerare è quando lo stato dell'automa descrive un binario come vuoto, ma l'azione letta descrive l'uscita di un treno da tale binario. È una situazione impossibile da realizzarsi, e come tale può essere gestita tramite lo stato *errore*. Tuttavia, un modello più sofisticato potrebbe utilizzare uno stato aggiuntivo per catturare le situazioni fisicamente impossibili.

Infine, è evidente che se un treno entrante è deviato su un binario occupato si incorre in un malfunzionamento del sistema stazione. Questa situazione, nel modello, dovraà portare l'automa nello stato *errore*.

L'insieme dei simboli, Σ, può essere:

$$\Sigma = \{i,\, s,\, u_a,\, u_b,\, t\}$$

dove i rappresenta l'ingresso di un treno dall'esterno e l'occupazione della linea esterna a monte dello scambio, s l'azionamento dello scambio, u_a e u_b rappresentano l'uscita di un treno dai binari a e b, rispettivamente, e t desscrive il trascorrere del tempo (abbastanza perchè un treno transiti sui tratti di binario considerati nel modello).

Ogni sequenza di azioni che non comporti il raggiungimento dello stato *errore* rappresenta il normale funzionamento della stazione. 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 stazione libera e lo scambio diretto sul binario a: $la00$.

Con le ipotesi fatte, dovrebbero essere accettate, per esempio, le seguenti sequenze di azioni: $iisi$, $istu_b$. Al contrario, verrebbero rifiutate, tra le altre, le seguenti sequenze di azioni: u_a, iii, $siiu_b$. 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, $\delta : Q \times \Sigma \rightarrow Q$ può essere quella riportata in Tabella 1.

Alcune varianti sono possibili; tra le altre:

- i treni possono/devono fermarsi (altre azioni e molte possibilità di descrivere situazioni critiche);
- si può ipotizzare che lo scambio si porti automaticamente sul primo binario libero (il numero di stati si dimezza);
- si può ipotizzare che i treni escano in ordine di binario (si elimina un'azione).

Per meglio comprenderlo, il modello può essere semplificato eliminando un bianrio (e lo scambio).

Esercizio 5

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

- $E = (ba^* + ac)^2(b^3 + c^*ab)^*$

Individuare, motivando le risposte, quali fra le seguenti stringhe vengono descritte da E:

a) *baaacbbb*

b) *acaccabccab*

c) *ababbbb*

d) *cabbca*

e) *bbabbc*

f) *ccacaab*

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

$\boldsymbol{\delta}$	i	s	u_a	u_b	t
$la00$	$ia00$	$lb00$	*errore*	*errore*	$la00$
$la01$	$ia01$	$lb01$	*errore*	$ua00$	$la01$
$la10$	$ia10$	$lb10$	$ua00$	*errore*	$la10$
$la11$	$ia11$	$lb11$	$ua01$	$ua10$	$la11$
$lb00$	$ib00$	$la00$	*errore*	*errore*	$lb00$
$lb01$	$ib01$	$la01$	*errore*	$ub00$	$lb01$
$lb10$	$ib10$	$la10$	$ub00$	*errore*	$lb10$
$lb11$	$ib11$	$la11$	$ub01$	$ub10$	$lb11$
$ia00$	$ia10$	$ib00$	*errore*	*errore*	$la10$
$ia01$	$ia11$	$ib01$	*errore*	*errore*	$la11$
$ia10$	*errore*	$ib10$	*errore*	*errore*	*errore*
$ia11$	*errore*	$ib11$	*errore*	*errore*	*errore*
$ib00$	$ib01$	$ia00$	*errore*	*errore*	$lb01$
$ib01$	*errore*	$ia01$	*errore*	*errore*	*errore*
$ib10$	$ib11$	$ia10$	*errore*	*errore*	$lb11$
$ib11$	*errore*	$ia11$	*errore*	*errore*	*errore*
$ua00$	*errore*	$ub00$	*errore*	*errore*	$ua00$
$ua01$	*errore*	$ub01$	*errore*	$ua00$	$la01$
$ua10$	*errore*	$ub10$	$ua00$	*errore*	$la10$
$ua11$	*errore*	$ub11$	$ua01$	$ua10$	$la11$
$ub00$	*errore*	$ua00$	*errore*	*errore*	$la00$
$ub01$	*errore*	$ua01$	*errore*	$ub00$	$lb01$
$ub10$	*errore*	$ua10$	$ub00$	*errore*	$lb10$
$ub11$	*errore*	$ua11$	$ub01$	$ub10$	$lb11$
errore	*errore*	*errore*	*errore*	*errore*	*errore*

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

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: $E_1 = (ba^* + ac)^2$ e $E_2 = (b^3 + c^*ab)^*$. Quindi, le stringhe descritte da E dovranno obbligatoriamente avere un prefisso descritto da E_1 eventualmente seguito da un suffisso descritto da E_2 (poiché E_2 descrive anche la stringa vuota). Questa premessa semplificherà la spiegazione delle soluzioni di seguito riportate.

- *baaacbbb*
  $baaacbbb = (baa)(ac)(bbb) \subseteq (ba^2)(ac)(b^3) \subseteq (ba^*)(ac)(b^3 + c^*ab) \subseteq (ba^* + ac)^2(b^3 + c^*ab)^*$

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

  Spiegazione alternativa:

  $b \quad a \quad a \quad a \quad c \quad b \quad b \quad b$
  $baa \qquad ac \qquad bbb$
  $ba^* \qquad ac \qquad b^3$
  $(ba^* + ac)^2 \qquad (b^3 + c^*ab)$
  $(ba^* + ac)^2 \qquad (b^3 + c^*ab)^*$

- *acaccabccab*
  $acaccabccab = (ac)(ac)(cab)(ccab) \subseteq (ba^* + ac)(ba^* + ac)(c^*ab)(c^*ab) \subseteq (ba^* + ac)^2(b^3 + c^*ab)(b^3 + c^*ab) \subseteq (ba^* + ac)^2(b^3 + c^*ab)^*$

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

  Spiegazione alternativa:

  $a \quad c \quad a \quad c \quad c \quad a \quad b \quad c \quad c \quad a \quad b$
  $ac \qquad ac \qquad cab \qquad ccab$
  $(ac)^2 \qquad c^*ab \qquad c^*ab$
  $(ba^* + ac)^2 \qquad (c^*ab)^*$
  $(ba^* + ac)^2 \qquad (b^3 + c^*ab)^*$

- *ababbbb*
  Le stringhe descritte da E non possono iniziare per *ab*.

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

- *cabbca*
  Le stringhe descritte da E non possono iniziare per *c*.

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

- *bbabbc*
  Il suffisso *bba* (o anche *bb*) potrebbe essere descritto da E_1, ma la rimanente sottostringa, *bbc* (o *abbc*), non può essere descritta da E_2.

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

- *ccacaab*
  Le stringhe descritte da E non possono iniziare per c.

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

Esercizio 6

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

- *baaabaca*
- *bbcacaca*
- *baabcbaa*
- *baabacbcb*

ma non le seguenti:

- *bbbaca*
- *abcbba*
- *abbccc*
- *cabba*

Soluzione

Si può notare che i primi tre simboli sono sufficienti per distinguere le stringhe del primo insieme da quelle del secondo insieme. Infatti, le stringhe del primo gruppo hanno come prefisso *baa* o *bbc*. L'espressione regolare che descrive questa caratteristica è $b(a^2+bc)(a+b+c)^*$. Nessuna delle stringhe del secondo gruppo viene descritta da tale espressione regolare:

- *bbbaca*: inizia per *bbb*;
- *abcbba*: non inizia per b;
- *abbccc*: non inizia per b;
- *cabba*: non inizia per b.

Altre espressioni regolari che rispettano le specifiche:

- $ba^*(bc+ba)(a+b+c)^*$
- $b(a^2+bc)(a+ba^*)(ca+ba+a^*+cb)^2$
- $(ba^*)^2a^*(cba^*+ca)^*$
- $b(a+b)(a+c)(a+b+c)^*$
- $b(bc+aa)(a+b+c)^*$
- $b(a^2+bc)(a+b+c)^*$
- $(ba^*+bc)^2(ac^*+cb+cba^*)^*$

Una descrizione alternativa può essere ricavata notando che tutte le stringhe da includere sono lunghe almeno 8 simboli e non più di 9, mentre le stringhe da escludere hanno, al più, lunghezza 6. Pertanto, l'espressione $(a+b+c)^8+(a+b+c)^9$ soddisfa i requisiti. Analogo ragionamento può essere fatto per motivare l'espressione regolare $(a+b+c)^8b^*$.