

Nome, Cognome: \_\_\_\_\_ Numero matricola: \_\_\_\_\_

# Fondamenti di Informatica II

## I prova intermedia

1

23 Maggio 2003

1. Segnare con una "X" le affermazioni corrette

**a) Analisi lessicale**

- ☐ L'analizzatore lessicale analizza un programma sorgente e produce l'albero sintattico
- ☐ L'operatore "|" delle espressioni regolari **non** soddisfa la proprietà associativa
- ☐ L'espressione regolare  $aa^*(b|c)$  è equivalente a  $(a^*ab|a^*ab)$

**b) Analisi sintattica**

- ☐ YACC è uno strumento che permette di generare un analizzatore sintattico a partire dalla grammatica del linguaggio
- ☐ Un analizzatore discendente partendo dallo scopo  $S$  tenta di ricostruire l'albero di derivazione con cui si ottiene la stringa  $w$  da riconoscere
- ☐ Una grammatica in cui sia stata eliminata la ricorsione a sinistra è sicuramente non ambigua

**c) Analisi sintattica ascendente**

- ☐ Sia  $uww$  una stringa di terminali e non terminali e  $A \rightarrow w$  una produzione del linguaggio  $L(G)$ . Si supponga che un analizzatore ascendente LR (correttamente realizzato per il linguaggio  $L(G)$ ) sostituisca  $w$  con  $A$ , riducendo  $uww$  a  $uAv$ . In tal caso, se  $uww$  appartiene al linguaggio  $L(G)$  deve esistere una derivazione destra dallo scopo  $S$  a  $uww$ .
- ☐ La tabella di analisi di un analizzatore LR è uguale a quella di un analizzatore discendente: contiene una sola tabella in cui per ogni ingresso e ogni simbolo non terminale si indica con quale produzione espandere.

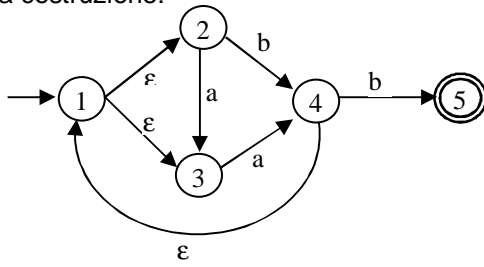
**d) Grammatiche e linguaggi**

- ☐ Se un linguaggio è  $LL(1)$  allora è anche  $LL(k)$  per ogni  $k > 1$ .
- ☐ Una grammatica che contiene una produzione  $vAw \rightarrow vuw$  dove  $A$  è un non simbolo non terminale,  $v, w, u$  stringhe di non terminali e terminale è una grammatica non contestuale.

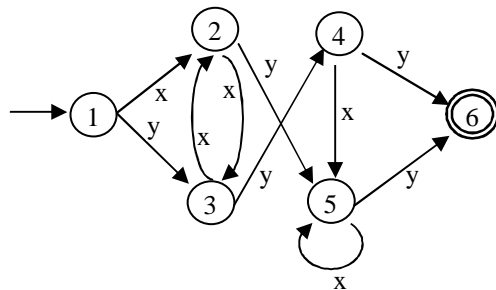
2. Si disegni l'automa a stati finiti che riconosce la seguente espressione regolare:

**$a^*c(b|c)^*a$**

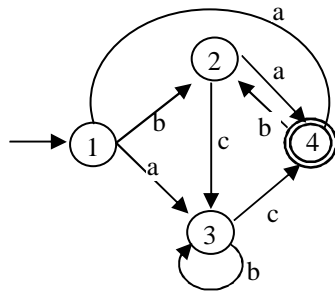
3. Dato il seguente automa a stati finiti si disegni l'automa deterministico equivalente mostrando come si è giunti alla sua costruzione:



4. Dato il seguente automa a stati finiti si costruisca l'automa minimo equivalente mostrando i passi intermedi:



5. Dato il seguente automa a stati si trovi l'espressione regolare che esso riconosce



6. Il comando Unix “ls -l” mostra il contenuto di una directory. Per ogni file, vengono visualizzati da destra a sinistra, divisi da caratteri di spaziatura: 10 caratteri che definiscono il tipo di file e i diritti di accesso; per le directory, il numero di file che contiene (o 1 per i file); il nome dell’utente proprietario; il gruppo a cui appartiene il file; la dimensione; mese, giorno e ora dell’ultima modifica, se fatta nell’anno corrente, oppure mese, giorno e anno; il nome del file.

```

-rw-r--r-- 1 pippo users 44691 Jun 10 16:58 Fondamenti4.tex
-rwxr-xr-x 1 pippo users 100 Jun 10 17:21 Fourpages
-rw-r--r-- 1 pippo users 19968 Jun 12 17:06 MPS2.doc
drwx----- 2 pippo users 2048 Jun 10 17:31 Mail
-rw-r--r-- 1 pippo users 22528 Jun 12 10:21 Preventivo.doc
drwxr-xr-x 2 pippo users 2048 Oct 17 2001 bibbiena
drwx----- 2 pippo users 2048 Feb 15 11:24 mail
drwxr-xr-x 2 pippo users 2048 Jun 27 2000 public_html
drwxr-xr-x 3 pippo users 2048 Feb 21 17:24 seminari
drwxr-xr-x 59 pippo users 2048 May 13 11:03 tex
-rwxr-xr-x 1 pippo users 53 Jun 10 17:29 twopage

```

Si fornisca un’espressione per grep (posix 1003.2) che consenta di individuare tutte le righe relative a directory che contengono più di 80 file e appartengono al gruppo “user” o al gruppo “root”

7. Data la seguente grammatica si provi che è ambigua.

$S \rightarrow cABc \mid cc$

$A \rightarrow abA \mid ab$

$B \rightarrow aB \mid bB \mid \varepsilon$

8. Si fattorizzi a sinistra e si rimuova la ricorsione immediata a sinistra nella seguente grammatica.

$S \rightarrow aBd \mid ac \mid b$

$B \rightarrow Bd \mid a \mid Ab$

$A \rightarrow c \mid S$

9. Data la seguente grammatica

$S \rightarrow AB \mid cAS \mid b$

$A \rightarrow aAa \mid \varepsilon$

$B \rightarrow b \mid \varepsilon$

a) Si calcoli First e Follow

b) Si costruisca la tabella di analisi per l'analizzatore discendente e si dica se la grammatica è LL(1)

## Regular expressions

The fundamental building blocks are the regular expressions that match a single character. Most characters, including all letters and digits, are regular expressions that match themselves. Any metacharacter with special meaning may be quoted by preceding it with a backslash. A list of characters enclosed by '[' and ']' matches any single character in that list; if the first character of the list is the caret '^', then it matches any character **not** in the list. For example, the regular expression '[0123456789]' matches any single digit. A range of ASCII characters may be specified by giving the first and last characters, separated by a hyphen. Finally, certain named classes of characters are predefined, as follows. Their interpretation depends on the LC\_CTYPE locale; the interpretation below is that of the POSIX locale, which is the default if no LC\_CTYPE locale is specified.

'[:alnum:]'	Any of '[:digit:]' or '[:alpha:]'
'[:alpha:]'	Any letter: a b c d e f g h i j k l m n o p q r s t u v w x y z, A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.
'[:blank:]'	Space or tab.
'[:cntrl:]'	Any character with octal codes 000 through 037, or DEL (octal code 177).
'[:digit:]'	Any one of 0 1 2 3 4 5 6 7 8 9.
'[:graph:]'	Anything that is not a '[:alnum:]' or '[:punct:]'.
'[:lower:]'	Any one of a b c d e f g h i j k l m n o p q r s t u v w x y z.
'[:print:]'	Any character from the '[:space:]' class, and any character that is <b>not</b> in the '[:graph:]' class.
'[:punct:]'	Any one of ! " # \$ % & ' ( ) * + , - . / : ; < = > ? @ [ \ ] ^ _ ` {   } ~.
'[:space:]'	Any one of CR FF HT NL VT SPACE.
'[:upper:]'	Any one of A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.
'[:xdigit:]'	Any one of a b c d e f A B C D E F 0 1 2 3 4 5 6 7 8 9.

For example, '[[:alnum:]]' means '[0-9A-Za-z]', except the latter form is dependent upon the ASCII character encoding, whereas the former is portable. (Note that the brackets in these class names are part of the symbolic names, and must be included in addition to the brackets delimiting the bracket list.) Most metacharacters lose their special meaning inside lists. To include a literal ']', place it first in the list. Similarly, to include a literal '^', place it anywhere but first. Finally, to include a literal '-', place it last.

The period '.' matches any single character. The symbol '\w' is a synonym for '[[:alnum:]]' and '\W' is a synonym for '[^[:alnum:]]'. The caret '^' and the dollar sign '\$' are metacharacters that respectively match the empty string at the beginning and end of a line. The symbols '<' and '>' respectively match the empty string at the beginning and end of a word. The symbol '\b' matches the empty string at the edge of a word, and '\B' matches the empty string provided it's not at the edge of a word.

A regular expression may be followed by one of several repetition operators:

'?'	The preceding item is optional and will be matched at most once.
'*'	The preceding item will be matched zero or more times.
'+'	The preceding item will be matched one or more times.
'{n}'	The preceding item is matched exactly <i>n</i> times.
'{n,}'	The preceding item is matched <i>n</i> or more times.
'{n,m}'	The preceding item is matched at least <i>n</i> times, but not more than <i>m</i> times.

Two regular expressions may be concatenated; the resulting regular expression matches any string formed by concatenating two substrings that respectively match the concatenated subexpressions.

Two regular expressions may be joined by the infix operator '|'; the resulting regular expression matches any string matching either subexpression.

Repetition takes precedence over concatenation, which in turn takes precedence over alternation. A whole subexpression may be enclosed in parentheses to override these precedence rules.

The backreference '\n', where *n* is a single digit, matches the substring previously matched by the *n*th parenthesized subexpression of the regular expression.

In basic regular expressions the metacharacters '?', '+', '{', '|', '(', and ')' lose their special meaning; instead use the backslashed versions '\?', '\+', '\{', '\|', '\(', and '\)'.

Nome Cognome: \_\_\_\_\_ Numero matricola: \_\_\_\_\_

# Fondamenti di Informatica II

## I prova intermedia

2

23 Maggio 2003

1. Segnare con una "X" le affermazioni corrette

**1. Analisi lessicale**

- ☐ Lex è uno strumento che permette di generare analizzatori lessicali
- ☐ L'operatore "|" delle espressioni regolari soddisfa la proprietà commutativa
- ☐ L'espressione regolare  $a(b|c)a$  è equivalente a  $(ab|ac)(a|a)$

**2. Analisi sintattica**

- ☐ L'analizzatore sintattico di un compilatore C produce in uscita il codice eseguibile
- ☐ Un analizzatore ascendente partendo dalla stringa in ingresso  $w$  tenta di ricostruire l'albero di derivazione usando le regole della grammatica all'indietro
- ☐ Una grammatica che sia stata fattorizzata a sinistra è sicuramente non ambigua

**3. Analisi sintattica ascendente**

- ☐ Sia  $uww$  una stringa di terminali e non terminali e  $A \rightarrow w$  una produzione del linguaggio  $L(G)$ . Si supponga che un analizzatore ascendente LR (correttamente realizzato per il linguaggio  $L(G)$ ) sostituisca  $w$  con  $A$ , riducendo  $uww$  a  $uAv$ . In tal caso,  $v$  contiene solo simboli terminali.
- ☐ Lo stack di un analizzatore LR contiene solo simboli della grammatica.

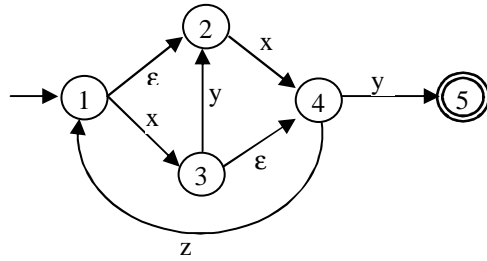
**4. Grammatiche e linguaggi**

- ☐ Se una grammatica è  $LL(1)$  allora è anche  $LR(1)$
- ☐ I linguaggi regolari sono contenuti nei linguaggi contestuali.

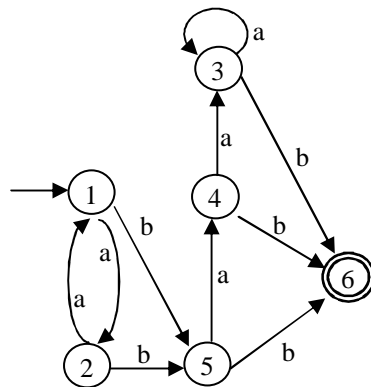
2. Si disegni l'automa a stati finiti che riconosce la seguente espressione regolare

$c(a^*|c)^*b$

3. Dato il seguente automa a stati finiti si disegni l'automa deterministico equivalente mostrando come si è giunti alla sua costruzione:

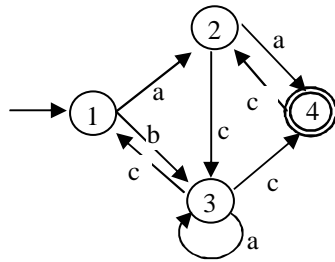


4. Dato il seguente automa a stati finiti si costruisca l'automa minimo equivalente mostrando i passi intermedi:





5. Dato il seguente automa a stati si trovi l'espressione regolare che esso riconosce



6. Il comando Unix “ps aux” mostra i processi attivi . Per ogni processo, vengono visualizzati da destra a sinistra, divisi da caratteri di spaziatura: il nome dell'utente proprietario del processo; un numero progressivo che identifica il processo; la percentuale del tempo di CPU usata; la percentuale di memoria usata; la dimensione virtuale (SIZE); la dimensione in memoria (RSS); il tty che controlla il processo; lo stato del processo; la data di inizio (o l'ora di inizio se la data è oggi); il tempo complessivo di CPU usato; il comando che ha iniziato il processo.

USER	PID	%CPU	%MEM	SIZE	RSS	TTY	STAT	START	TIME	COMMAND
at	198	0.0	0.0	892	112	?	S	Jun 12	0:00	/usr/sbin/atd
bin	131	0.0	0.1	820	212	?	S	Jun 12	0:04	/sbin/portmap
pippo	29455	0.0	1.0	1908	1280	p4	S	15:11	0:00	-bash
pippo	29500	0.0	3.5	6216	4524	p6	S	15:14	0:00	emacs surname.tex
pippo	30075	0.0	0.4	956	536	p4	R	18:13	0:00	ps aux
pluto	28271	0.0	1.0	1948	1324	p1	S	11:04	0:00	-bash
root	1	0.0	0.0	260	76	?	S	Jun 12	0:05	init [2]
root	4	0.0	0.0	0	0	?	SW	Jun 12	0:00	(kpiod)
root	5	0.0	0.0	0	0	?	SW	Jun 12	0:01	(kswapd)
root	74	0.0	0.2	1120	348	?	S	Jun 12	0:05	/usr/lib/postfix/mast

Si fornisca un'espressione per grep (posix 1003.2) che consenta di individuare tutte le righe relative a processi che: sono iniziati dai comandi “-bash” e “-tcsh”, sono in esecuzione da più di un ora e mezza e si trovano nello stato “S”.

7. Data la seguente grammatica si provi che è ambigua.

$S \rightarrow aSa \mid bSb \mid A$

$A \rightarrow aB \mid a$

$B \rightarrow aA \mid bB \mid \varepsilon$

8. Si fattorizzi a sinistra e si rimuova la ricorsione immediata a sinistra nella seguente grammatica.

$S \rightarrow ScDB \mid a \mid D$

$D \rightarrow dD \mid daD \mid d$

$B \rightarrow S \mid DS$

9. Data la seguente grammatica

$S \rightarrow cB \mid \varepsilon$

$A \rightarrow aA \mid \varepsilon$

$B \rightarrow AS \mid bBb$

e) Si calcoli First e Follow

f) Si costruisca la tabella di analisi per l'analizzatore discendente e si dica se la grammatica è LL(1)

## Regular expressions

The fundamental building blocks are the regular expressions that match a single character. Most characters, including all letters and digits, are regular expressions that match themselves. Any metacharacter with special meaning may be quoted by preceding it with a backslash. A list of characters enclosed by '[' and ']' matches any single character in that list; if the first character of the list is the caret '^', then it matches any character **not** in the list. For example, the regular expression '[0123456789]' matches any single digit. A range of ASCII characters may be specified by giving the first and last characters, separated by a hyphen. Finally, certain named classes of characters are predefined, as follows. Their interpretation depends on the LC\_CTYPE locale; the interpretation below is that of the POSIX locale, which is the default if no LC\_CTYPE locale is specified.

'[:alnum:]'	Any of '[:digit:]' or '[:alpha:]'
'[:alpha:]'	Any letter: a b c d e f g h i j k l m n o p q r s t u v w x y z, A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.
'[:blank:]'	Space or tab.
'[:cntrl:]'	Any character with octal codes 000 through 037, or DEL (octal code 177).
'[:digit:]'	Any one of 0 1 2 3 4 5 6 7 8 9.
'[:graph:]'	Anything that is not a '[:alnum:]' or '[:punct:]'.
'[:lower:]'	Any one of a b c d e f g h i j k l m n o p q r s t u v w x y z.
'[:print:]'	Any character from the '[:space:]' class, and any character that is <b>not</b> in the '[:graph:]' class.
'[:punct:]'	Any one of ! " # \$ % & ' ( ) * + , - . / : ; < = > ? @ [ \ ] ^ _ ` {   } ~.
'[:space:]'	Any one of CR FF HT NL VT SPACE.
'[:upper:]'	Any one of A B C D E F G H I J K L M N O P Q R S T U V W X Y Z.
'[:xdigit:]'	Any one of a b c d e f A B C D E F 0 1 2 3 4 5 6 7 8 9.

For example, '[[:alnum:]]' means '[0-9A-Za-z]', except the latter form is dependent upon the ASCII character encoding, whereas the former is portable. (Note that the brackets in these class names are part of the symbolic names, and must be included in addition to the brackets delimiting the bracket list.) Most metacharacters lose their special meaning inside lists. To include a literal ']', place it first in the list. Similarly, to include a literal '^', place it anywhere but first. Finally, to include a literal '-', place it last.

The period '.' matches any single character. The symbol '\w' is a synonym for '[[:alnum:]]' and '\W' is a synonym for '[^[:alnum:]]'. The caret '^' and the dollar sign '\$' are metacharacters that respectively match the empty string at the beginning and end of a line. The symbols '<' and '>' respectively match the empty string at the beginning and end of a word. The symbol '\b' matches the empty string at the edge of a word, and '\B' matches the empty string provided it's not at the edge of a word.

A regular expression may be followed by one of several repetition operators:

'?'	The preceding item is optional and will be matched at most once.
'*'	The preceding item will be matched zero or more times.
'+'	The preceding item will be matched one or more times.
'{n}'	The preceding item is matched exactly <i>n</i> times.
'{n,}'	The preceding item is matched <i>n</i> or more times.
'{n,m}'	The preceding item is matched at least <i>n</i> times, but not more than <i>m</i> times.

Two regular expressions may be concatenated; the resulting regular expression matches any string formed by concatenating two substrings that respectively match the concatenated subexpressions.

Two regular expressions may be joined by the infix operator '|'; the resulting regular expression matches any string matching either subexpression.

Repetition takes precedence over concatenation, which in turn takes precedence over alternation. A whole subexpression may be enclosed in parentheses to override these precedence rules.

The backreference '\n', where *n* is a single digit, matches the substring previously matched by the *n*th parenthesized subexpression of the regular expression.

In basic regular expressions the metacharacters '?', '+', '{', '|', '(', and ')' lose their special meaning; instead use the backslashed versions '\?', '\+', '\{', '\|', '\(', and '\)'.