

ESERCIZIO
(tratto dalla prova in itinere 2012-13)

Con l'ausilio di Flex scrivere un programma che crei la tabella del lessico di un programma scritto nel linguaggio Abis. Tale linguaggio è case-sensitive e usa le seguenti parole chiave: **begin, end, do, while, if, then**. Le costanti sono numeri interi non negativi. L'attributo è il valore numerico corrispondente. Gli identificatori sono parole costituite da lettere e cifre e che cominciano per lettera. Alle seguenti tre categorie di identificatori il programma è in grado di assegnare un attributo:

- a. Identificatori di tipo 1: sono gli identificatori costituiti da lettere maiuscole tranne la prima e l'ultima che sono minuscole. L'attributo è il numero delle lettere maiuscole. Per esempio l'identificatore aAGTYEj ha attributo 6.
- b. Identificatori di tipo 2: sono gli identificatori che iniziano e finiscono per lettera maiuscola, e contenenti in mezzo solo cifre. L'attributo è il numero più alto di 0 consecutivi. Per esempio l'identificatore B2003400012S ha attributo 3.
- c. Identificatori di tipo 3: sono gli identificatori che iniziano e finiscono per consonante maiuscola e che contengono solo lettere al loro interno. L'attributo è il numero di vocali minuscole. Per esempio l'identificatore GsKietrdoguR ha attributo 4.

Per gli altri identificatori l'attributo è ND (non determinato). Se un identificatore o una costante si trovano dentro un costrutto do...while l'attributo raddoppia. Il programma non deve entrare nel merito della correttezza sintattica del costrutto. Si suppone pertanto che nel file in input non ci siano costrutti do...while annidati. Il programma, preso in input un file contenente codice scritto in Abis, deve restituire un file di output contenente una tabella con le informazioni sui token, sui lessemi, sugli attributi e sul numero di riga in cui ciascun token è presente. Alla fine della scansione il programma deve terminare con un messaggio: PROGRAMMA ANALIZZATO CORRETTAMENTE. Per esempio se il file in input è:

```
begin
aAGTYEj
do
B2003400012S    GsKietrdoguR
while
aghdji9A 65700
end
```

il file in output è:

TOKEN	LESSEMA	ATTRIBUTO	RIGA
Keyword	begin		1
ID1	aAGTYEj	6	2
Keyword	do		3
ID2	B2003400012S	6	4
ID3	GsKietrdoguR	8	4
Keyword	while		5
ID	aghdji9A	ND	6
CONST	65700	65700	6
Keyword	end		7

PROGRAMMA ANALIZZATO CORRETTAMENTE