

A dark blue vertical bar on the left side of the page. A blue arrow points to the right from the bar, containing the text 'ΕΡΓΑΣΤΗΡΙΟ'.

ΕΡΓΑΣΤΗΡΙΟ

# ΜΕΤΑΓΛΩΤΤΙΣΤΕΣ

ΜΕΡΟΣ Α2

Several thin, curved lines in dark blue and light grey originate from the bottom left and curve upwards and to the right.

## ΤΜΗΜΑ Β1

ΠΕΤΡΑΚΗ ΒΑΣΙΛΙΚΗ ΕΥΑΝΘΙΑ 19390193  
ΠΑΠΑΧΡΙΣΤΟΔΟΥΛΟΥ ΑΙΚΑΤΕΡΙΝΗ 19390185  
ΡΟΥΜΕΛΙΩΤΗΣ ΣΠΥΡΙΔΩΝ 19390205  
ΜΑΡΓΙΩΛΗΣ ΧΡΗΣΤΟΣ 19390133

## Περιεχόμενα

Στόχος εργασίας.....	3
Κανονικές εκφράσεις .....	4
Λεκτικές μονάδες.....	4
Αυτόματα .....	5
Κώδικας FSM.....	5
Ακέρατοι.....	5
Αριθμοί Κινητής Υποδιαστολής .....	6
Ονόματα ορισμών και στοιχείων γεγονότων .....	7
Μεταβλητές .....	7
Συμβολοσειρές.....	8
Σχόλια.....	8
Διαχωριστές .....	9
Πίνακες Μετάβασης .....	10
Μεταβλητές .....	10
Διαχωριστές .....	10
Ονόματα ορισμών και στοιχείων γεγονότων .....	10
Σχόλια.....	11
Αριθμοί Κινητής Υποδιαστολής .....	11
Ακέρατοι Αριθμοί .....	11
Συμβολοσειρές.....	11
Ενιαίο .....	12
Πίνακας Μετάβασης.....	12
.....	12
Ενιαίο FSM .....	13
Ενδεικτικά Τρεξίματα Ενιαίου .....	15
Αποδεκτές Είσοδοι.....	15
Μη αποδεκτές είσοδοι .....	17
Ανάλυση Αρμοδιοτήτων .....	18

Βιβλιογραφία .....	19
--------------------	----

## Στόχος εργασίας

---

Στην παρακάτω εργασία, ως ομάδα, αναλύουμε ένα εισαγωγικό κομμάτι των μεταγλωττιστών, συγκεκριμένα τις κανονικές εκφράσεις και τα πεπερασμένα αυτόματα. Αυτό θα γίνει πάνω στην εικονική γλώσσα προγραμματισμού «Uni-CLIPS» και χρησιμοποιούμε το εργαλείο FSM. Για την απεικόνιση των αυτόματων θα χρησιμοποιήσουμε το εργαλείο Microsoft Visio.

- **Ακέρατοι αριθμοί**

1. `[+-]?[1-9]{1}[0-9]*`

- Αριθμοί κινητής υποδιαστολής

1. `[+-]?[0-9]+((\.[0-9]+)([eE][+-]?[0-9]*)?)|([eE][+-]?[0-9]*)?)`

- **Ονόματα ορισμών και στοιχείων γεγονότων**

1.  $[A - Z a - z] + [A - Z a - z \theta - 9 \_ -]^*$

- **Μεταβλητές**

1. `\?[A-Za-z0-9]+`

- Συμβολοσειρές

1. `"[^"\\]*(?:\\.["\\]*)*"`

- **Σχόλια**

- Διαχωριστές

1. `[\n\ \t]+`

# Αυτόματα

Κώδικας FSM

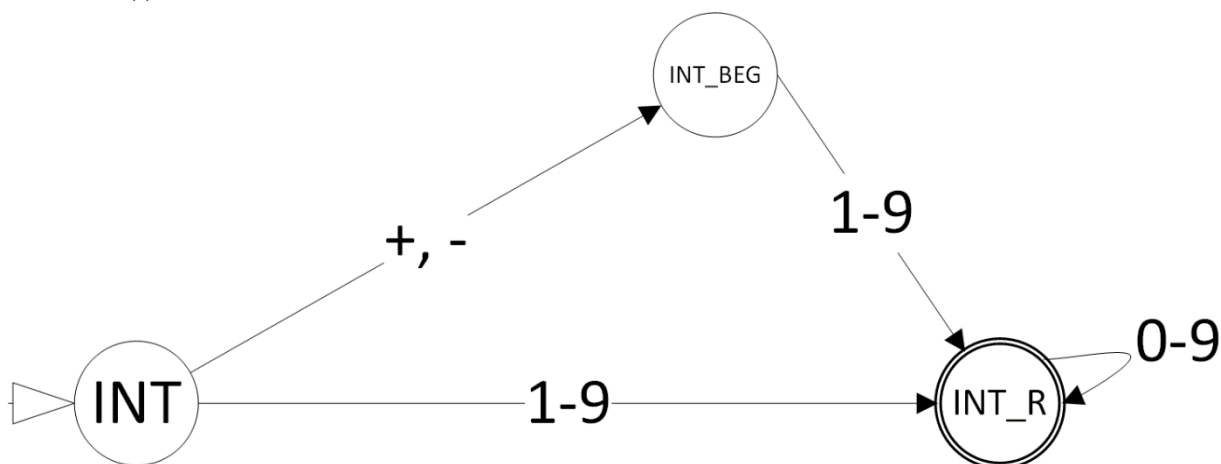
Ακέραιοι

```

1. START=INT
2.
3. // Έχουμε 2 περιπτώσεις, στην 1 αρχίζει με +/- και στην 2η με 1-9
4. INT:      + \-      -> INT_BEG // 1η περίπτωση
5.           1-9       -> INT_R  // Δεύτερη περίπτωση
6.           *         -> BAD  // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
7. // Μετά το +/-, συνεχίζει με ψηφίο 1-9
8. INT_BEG:  1-9       -> INT_R
9.           *         -> BAD  // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
10. // Μετά το 1-9, συνεχίζει με 0-9
11. INT_R:   0-9       -> INT_R
12.          \n        -> GOOD
13.          *         -> BAD  // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
14. GOOD(OK): \n      -> GOOD
15.

```

Απεικόνιση με Visio



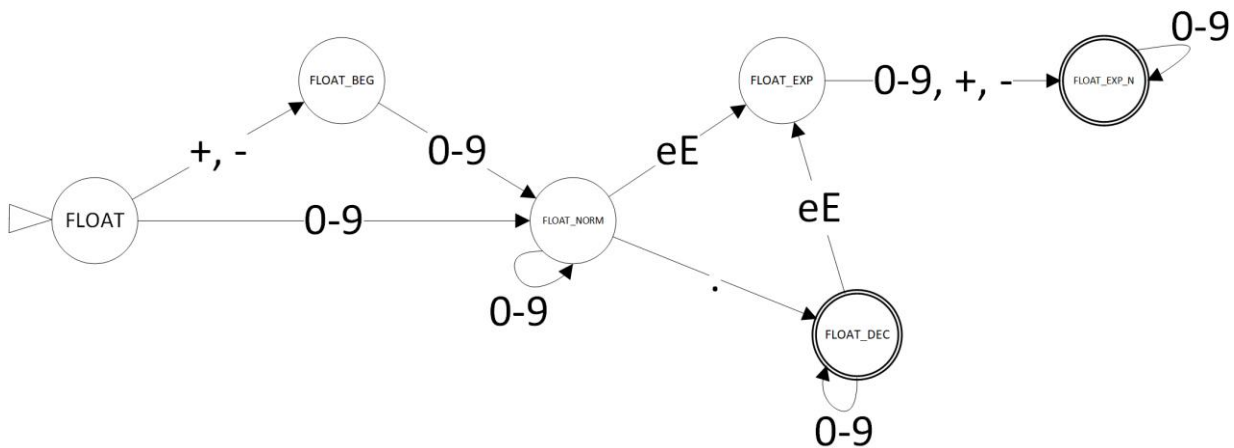
## Αριθμοί Κινητής Υποδιαστολής

```

1.  START=FLOAT
2.
3.  // Έχουμε τις περιπτώσεις όπου δίνεται +/- ή 0-9
4.  FLOAT:      + \-          -> FLOAT_BEG
5.              0-9          -> FLOAT_NORM
6.              *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
7.  // Για +/-
8.  FLOAT_BEG:  0-9          -> FLOAT_NORM
9.              *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
10. // Για 0-9
11. FLOAT_NORM: 0-9          -> FLOAT_NORM // Loop
12.             e E          -> FLOAT_EXP // Εάν υπάρξει δύναμη
13.             .            -> FLOAT_DEC // Δεκαδικό
14.             *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
15. FLOAT_DEC:  0-9          -> FLOAT_DEC
16.             e E          -> FLOAT_EXP
17.             \n           -> GOOD
18.             *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
19. // Εάν ο αριθμός συνεχιστεί με τελεία
20. FLOAT_EXP:  0-9 + \-     -> FLOAT_EXP_N
21.             *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
22. FLOAT_EXP_N: 0-9        -> FLOAT_EXP_N
23.             \n           -> GOOD
24.             *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισοδοι
25. GOOD(OK):   \n           -> GOOD
26.

```

## Απεικόνιση με Visio



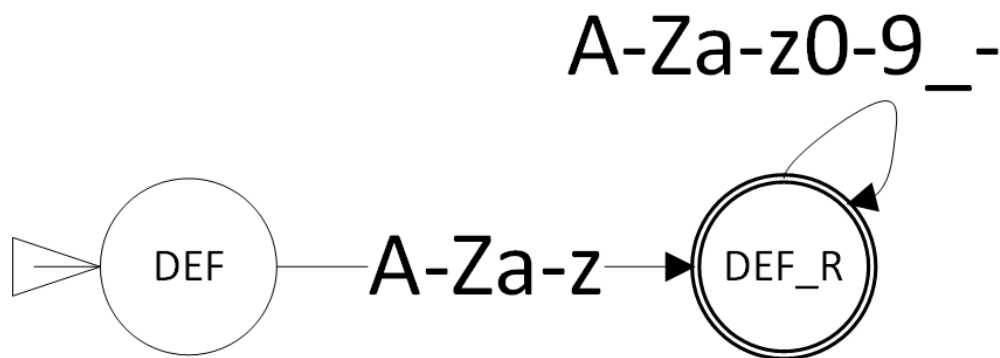
Ονόματα ορισμών και στοιχείων γεγονότων

```

1.  START=DEF
2.
3.  // Ορισμός μπορεί να ξεκινήσει με λατινικούς χαρακτήρες κεφαλαίους ή πεζούς
4.  DEF:      A-Z a-z          -> DEF_R
5.          *                  -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισόδους
6.  // Ύστερα συνεχίζει με λατινικούς χαρακτήρες, ψηφία 0-9, παύλα, κάτω παύλα και \
7.  DEF_R:    A-Z a-z 0-9 _ \- -> DEF_R
8.          *                  -> GOOD // Οποιοσδήποτε από τους χαρακτήρες εισόδου
9.  GOOD(OK): \n -> GOOD
10.

```

Απεικόνιση με Visio



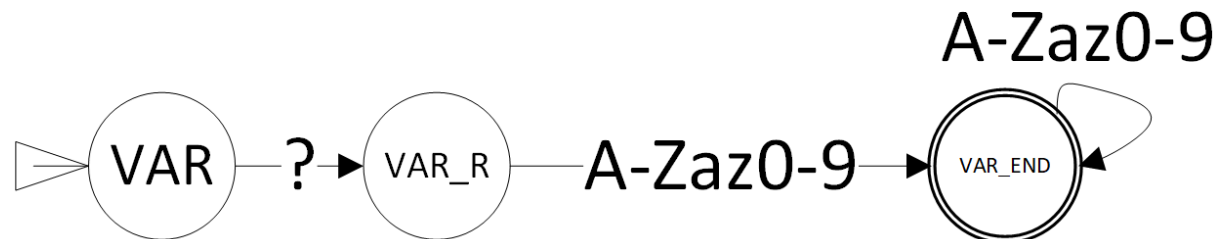
Μεταβλητές

```

1.  START=VAR
2.
3.  // Οι μεταβλητές πρέπει να ξεκινούν μόνο με ?
4.  VAR:      \?              -> VAR_R
5.          *                  -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισόδους
6.  // Στην συνέχεια είναι δεκτοί οι λατινικοί χαρακτήρες (πεζοί και κεφαλαίοι) και τα ψηφία 0-9
7.  VAR_R:    A-Z a-z 0-9     -> VAR_R
8.          \n                -> GOOD
9.          *                  -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισόδους
10. GOOD(OK): \n            -> GOOD
11.

```

Απεικόνιση με Visio





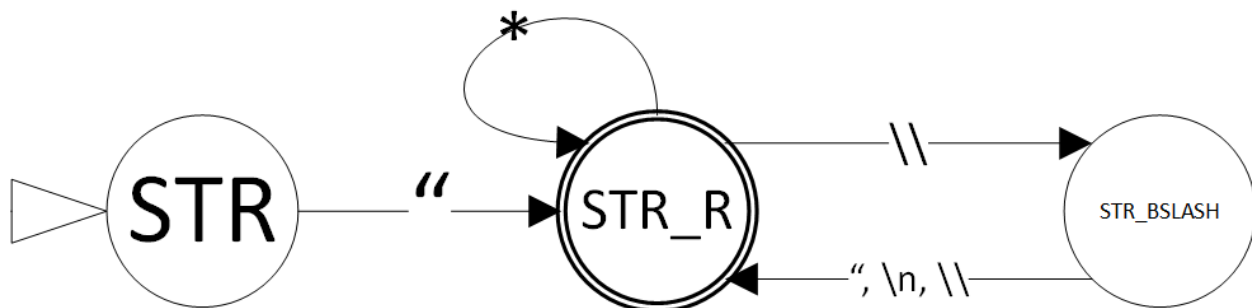
## Συμβολοσειρές

```

1.  START=STR
2.  // Δέχεται ", ή \
3.  STR:      "      -> STR_R
4.          *      -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισόδους
5.  STR_R:    "\     -> STR_BSLASH
6.          "     -> GOOD
7.          *     -> STR_R // Οποιοσδήποτε από τους χαρακτήρες εισόδου
8.  STR_BSLASH: " n \ -> STR_R
9.          *     -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισόδους
10.
11. GOOD(OK): \n     -> GOOD
12.

```

## Απεικόνιση με Visio



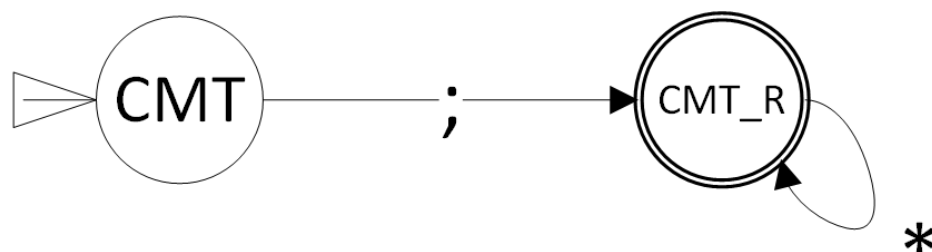
## Σχόλια

```

1.  START=CMT
2.
3.  // Ξεκινάει με ;
4.  CMT:      ;      -> CMT_R
5.          *      -> BAD
6.  CMT_R:    \n     -> GOOD
7.          *      -> CMT_R // Οποιοσδήποτε από τους χαρακτήρες εισόδου
8.  GOOD(OK) :
9.

```

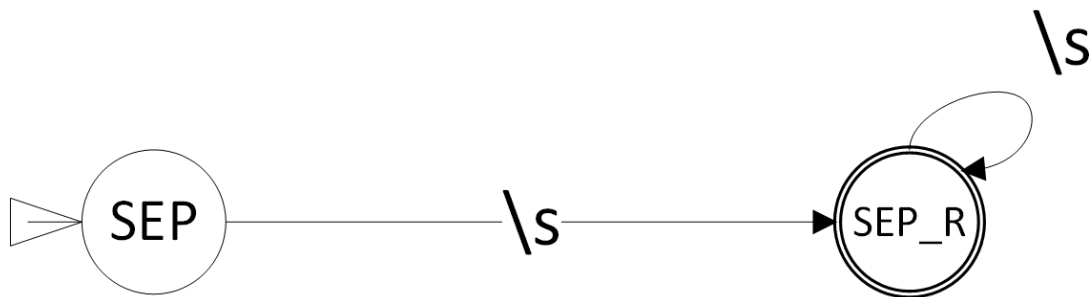
## Απεικόνιση με Visio



## Διαχωριστές

```
1. START=SEP
2.
3. // Οποιοσδήποτε white space χαρακτήρας
4. SEP:      \s          -> SEP_R
5.          *            -> BAD // Απορρίπτονται όλες οι υπόλοιπες εισόδους
6. SEP_R:    \s          -> SEP_R
7.          *            -> GOOD // Οποιοσδήποτε από τους χαρακτήρες εισόδου
8. GOOD(OK) :
9.
```

Απεικόνιση με Visio



## Πίνακες Μετάβασης

### Μεταβλητές

Χαρακτήρες Εισόδου						
Καταστάσεις		\?	A-Z a-z	0-9	*	\n
	START	VAR	GOOD	GOOD	BAD	BAD
	VAR	VAR_R	GOOD	GOOD	BAD	GOOD
	VAR_R	BAD	BAD	BAD	BAD	GOOD
	GOOD	BAD	BAD	BAD	BAD	GOOD

### Διαχωριστές

Χαρακτήρες Εισόδου			
Καταστάσεις		\s	*
	SEP	SEP_R	BAD
	SEP_R	SEP_R	GOOD

### Ονόματα ορισμών και στοιχείων γεγονότων

Χαρακτήρες Εισόδου								
Καταστάσεις		A-Z	a-z	*	0-9	\-	_	\n
	DEF	DEF_R	DEF_R	BAD	BAD	BAD	BAD	BAD
	DEF_R	GOOD	DEF_R	GOOD	DEF_R	DEF_R	DEF_R	GOOD
	GOOD	BAD	BAD	BAD	BAD	BAD	BAD	GOOD

### Σχόλια

Χαρακτήρες Εισόδου				
Καταστάσεις		;	\n	*
	CMT	CMT_R	BAD	BAD
	CMT_R	BAD	GOOD	GOOD

### Αριθμοί Κινητής Υποδιαστολής

Χαρακτήρες Εισόδου							
Καταστάσεις		+ -	0-9	e E	.	*	\n
	FLOAT	FLOAT_BEG	FLOAT_NORM	BAD	BAD	BAD	BAD
	FLOAT_BEG	BAD	FLOAT_NORM	BAD	BAD	BAD	BAD
	FLOAT_NORM	BAD	FLOAT_NORM	FLOAT_EXP	FLOAT_DEC	BAD	BAD
	FLOAT_DEC	BAD	FLOAT_DEC	FLOAT_EXP	BAD	BAD	GOOD
	FLOAT_EXP	FLOAT_EXP_N	FLOAT_EXP_N	BAD	BAD	BAD	GOOD
	FLOAT_EXP_N	BAD	FLOAT_EXP_N	BAD	BAD	BAD	GOOD
	GOOD	BAD	BAD	BAD	BAD	BAD	GOOD

### Ακέραιοι Αριθμοί

Χαρακτήρες Εισόδου						
Καταστάσεις		+ -	1-9	0-9	\n	*
	INT	INT_BEG	INT_R	BAD	BAD	BAD
	INT_BEG	BAD	INT_R	BAD	BAD	BAD
	INT_R	BAD	BAD	INT_R	GOOD	BAD
	GOOD	BAD	BAD	BAD	GOOD	BAD

### Συμβολοσειρές

Χαρακτήρες Εισόδου					
Καταστάσεις		\\	“	*	\n
	STR	STR_BACKSLASH	STR_R	BAD	BAD
	STR_R	STR_BSLASH	GOOD	STR_R	BAD
	STR_BSLASH	STR_R	STR_R	BAD	STR_R
	GOOD	BAD	BAD	BAD	GOOD

Ενιαίο

Πίνακας Μετάβασης

	+	\-	0	1-9	\?	A-Z	a-z	;	\s	"	0-9	_	.	e	E	\\	"	*	n	\n
SZ	NUM	NUM	Float_N ORM	INT	VAR	DEF	DEF	CMT	SEP	STR										
NUM			Float_N ORM	INT																
INT											INT		Float_DEC							GOOD
Float_N ORM											Float_NORM			Float_EXP	Float_EXP					GOOD
Float_D EC											Float_DEC			Float_EXP_N	Float_EXP_N					GOOD
Float_E XP	Float_E XP_N	Float_E XP_N									Float_EXP_N									
Float_E XP_N											Float_EXP_N									GOOD
STR																STR_Blash STR	GOOD STR	STR		
STR_Bla SH																			STR	
DEF		DEF				DEF	DEF				DEF	DEF						GOOD		
VAR						VAR	VAR				VAR									
VAR_R						VAR_R	VAR_R				VAR_R									GOOD
CMT																		CMT		GOOD
SEP									SEP									GOOD		
GOOD																				GOOD

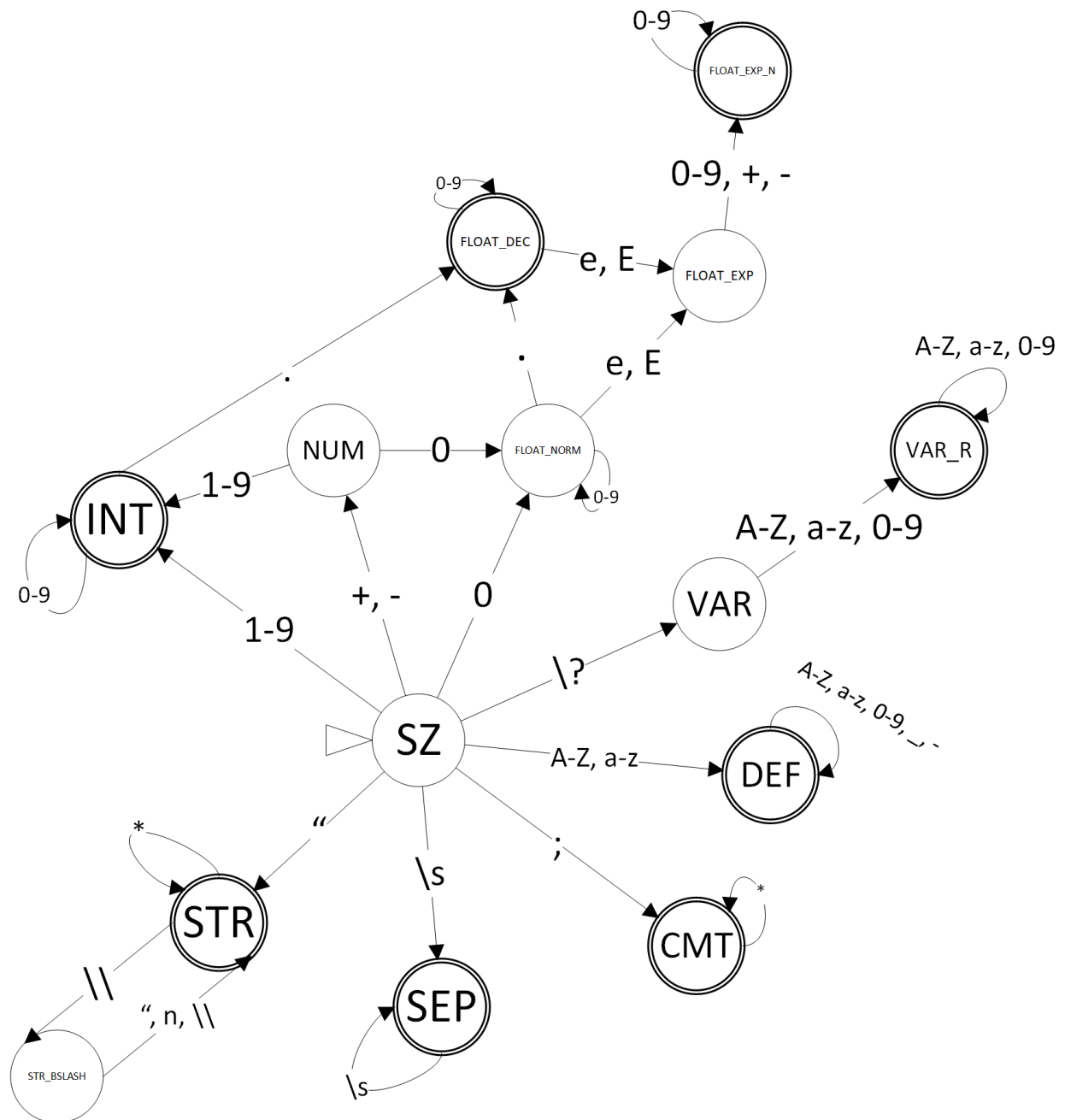
## Ενταίο FSM

```

1.  START=SZ
2.
3.  SZ:                                + \-      -> NUM
4.                                0              -> FLOAT_NORM
5.                                1-9            -> INT
6.                                \?            -> VAR
7.                                A-Z a-z      -> DEF
8.                                ;              -> CMT
9.                                \s            -> SEP
10.                               "              -> STR
11.
12.  NUM:                               0          -> FLOAT_NORM
13.                               1-9            -> INT
14.                               *              -> BAD
15.
16.  INT:                               0-9          -> INT
17.                               .              -> FLOAT_DEC
18.                               \n            -> GOOD
19.                               *              -> BAD
20.
21.  FLOAT_NORM: 0-9                  -> FLOAT_NORM
22.                               e E            -> FLOAT_EXP
23.                               .              -> FLOAT_DEC
24.                               *              -> BAD
25.  FLOAT_DEC:  0-9                  -> FLOAT_DEC
26.                               e E            -> FLOAT_EXP
27.                               \n            -> GOOD
28.                               *              -> BAD
29.  FLOAT_EXP:  0-9 + \-             -> FLOAT_EXP_N
30.                               *              -> BAD
31.  FLOAT_EXP_N: 0-9                  -> FLOAT_EXP_N
32.                               \n            -> GOOD
33.                               *              -> BAD
34.
35.  STR:                               \\\          -> STR_BSLASH
36.                               "              -> GOOD
37.                               *              -> STR
38.  STR_BSLASH: " n \\\             -> STR
39.                               *              -> BAD
40.
41.  DEF:                               A-Z a-z 0-9 _ \- -> DEF
42.                               *              -> GOOD
43.
44.  VAR:                               A-Z a-z 0-9      -> VAR_R
45.                               *              -> BAD
46.  VAR_R:                               A-Z a-z 0-9      -> VAR_R
47.                               \n            -> GOOD
48.                               *              -> BAD
49.
50.  CMT:                               \n            -> GOOD
51.                               *              -> CMT
52.
53.  SEP:                               \s            -> SEP
54.                               *              -> GOOD
55.
56.  GOOD(OK):  \n                      -> GOOD
57.

```

Απεικόνιση με Visio



## Ενδεικτικά Τρεξίματα Ενιαίου

## Αποδεκτές Είσοδοι

```
1. $ ./fsm lex.txt
2. +1234
3. YES
4.
5. $ ./fsm lex.txt
6. 50
7. YES
8.
9. $ ./fsm lex.txt
10. -115
11. YES
12.
13. $ ./fsm lex.txt
14. 3.14
15. YES
16.
17. $ ./fsm lex.txt
18. -10.0
19. YES
20.
21. $ ./fsm lex.txt
22. +0.0001
23. YES
24.
25. $ ./fsm lex.txt
26. 3.14e-10
27. YES
28.
29. $ ./fsm lex.txt
30. 0e0
31. YES
32.
33. $ ./fsm lex.txt
34. static-facts
35. YES
36.
37. $ ./fsm lex.txt
38. MoveUp
39. YES
40.
41. $ ./fsm lex.txt
42. CUBES
43. YES
44.
45. $ ./fsm lex.txt
46. sum-1
47. YES
48.
49. $ ./fsm lex.txt
50. table
51. YES
52.
53. $ ./fsm lex.txt
54. pacman
55. YES
56.
57. $ ./fsm lex.txt
```



```
58. A-21-b
59. YES
60.
61. $ ./fsm lex.txt
62. ?x
63. YES
64.
65. $ ./fsm lex.txt
66. ?X
67. YES
68.
69. $ ./fsm lex.txt
70. ?3
71. YES
72.
73. $ ./fsm lex.txt
74. ?ad
75. YES
76.
77. $ ./fsm lex.txt
78. ?X1b23
79. YES
80.
81. $ ./fsm lex.txt
82. ?32AbC
83. YES
84.
85. $ ./fsm lex.txt
86. ?ABcd1234de
87. YES
88.
89. $ ./fsm lex.txt
90. ""
91. YES
92.
93. $ ./fsm lex.txt
94. "Test"
95. YES
96.
97. $ ./fsm lex.txt
98. "Hello world"
99. YES
100.
101. $ ./fsm lex.txt
102. "Mark said, \"Boo!\""
103. YES
104.
105. $ ./fsm lex.txt
106. ; this is a comment
107. YES
108.
109. $ ./fsm lex.txt
110.
111. YES
112.
```

## Μη αποδεκτές είσοδοι

```
1. $ ./fsm lex.txt
2. ? hello world
3. fsm: in lex.txt, state 'good' input h not accepted
4.
5. $ ./fsm lex.txt
6. 0123
7. NO
8.
9. $ ./fsm lex.txt
10. 1static-facts
11. fsm: in lex.txt, state 'bad' input t not accepted
12.
13. $ ./fsm lex.txt
14. static(facts
15. fsm: in lex.txt, state 'good' input f not accepted
16.
17. $ ./fsm lex.txt
18. "hello world
19. NO
20.
21. $ ./fsm lex.txt
22. hello ; comment
23. fsm: in lex.txt, state 'good' input ; not accepted
24.
```

## Ανάλυση Αρμοδιοτήτων

---

- Πετράκη Βασιλική Ευανθία

Συγγραφή παραδοτέου, Διαγράμματα (visio και πίνακες), FSM debugging

- Παπαχριστοδούλου Αικατερίνη

Διαγράμματα (visio και πίνακες), Regex, FSM Debugging

- Ρουμελιώτης Σπυρίδων

Regex, FSM Code

- Μαργιώλης Χρήστος

Regex, FSM Code

Η συγγραφή της εργασίας πραγματοποιούταν σε πραγματικό χρόνο, μέσω ομαδικής κλήσης, οπότε όλα τα μέλη είναι ενημερωμένα για όλα τα κομμάτια της.

## Βιβλιογραφία

---

- <https://stackoverflow.com/>
- <https://eclass.uniwa.gr/courses/CS118/>
- «Μεταγλωττιστές» - Alfred V. Aho, Monica S. Lam, Ravi Sethi, Jeffrey D. Ullman
- “Regular Expressions Cheat Sheet” by DaveChild