

Μεταγλωττιστές

Προγραμματιστική Εργασία #1



Αλέξανδρος Ζερβόπουλος  
ΑΜ: Π2015111

# 1 Αναγνώριση Ενδείξεων Ώρας με Ντετερμινιστικό Αυτόματο Πεπερασμένων Καταστάσεων (DFA)

Ορίζουμε το αυτόματο:

- $Q = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6\}$
- $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, :, .\}$
- Αρχική κατάσταση:  $\{q_0\}$
- $F = \{q_6\}$
- Συνάρτηση μετάβασης  $\delta$  όπως περιγράφεται από τον πίνακα μεταβάσεων (Table 1)

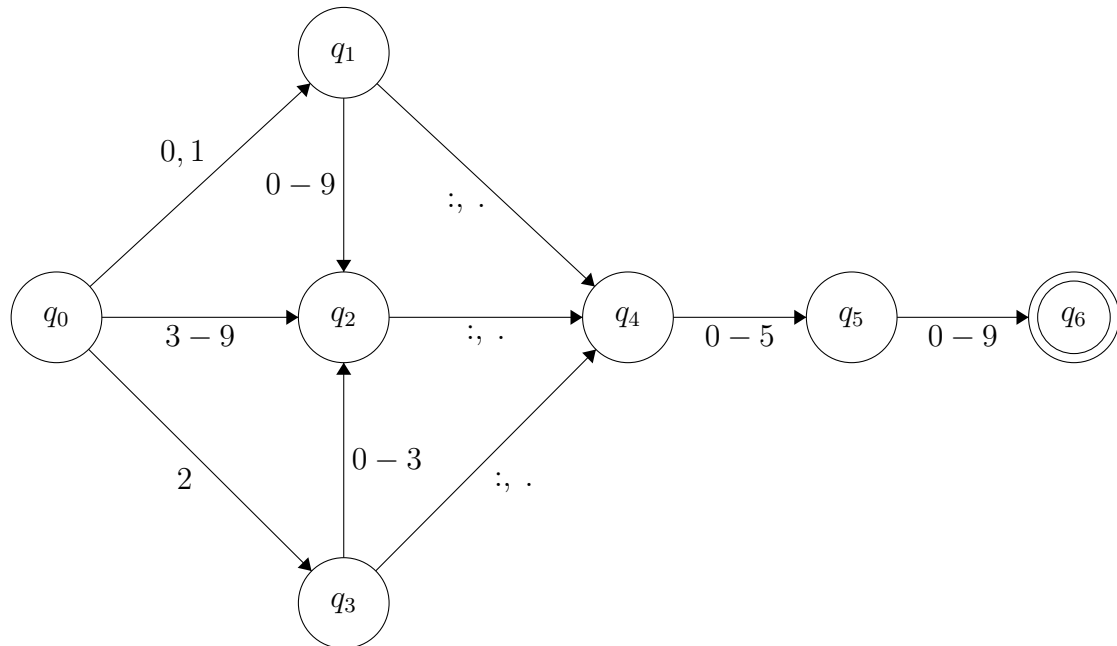
Το αυτόματο αυτό αναγνωρίζει ενδείξεις ώρας σε 24ωρη μορφή, αποτελούμενες από τα εξής μέρη:

1. Ώρα (0-23, σε 1 ή 2 ψηφία)
2. Διαχωριστής (: ή .)
3. Λεπτά (0-59, σε 2 ψηφία)

Πίνακας Μεταβάσεων						
	D_UPTO1	D_2	D_3	D_UPTO5	D_ANY	DELIMITER
$q_0$	$q_1$	$q_3$	$q_2$	$q_2$	$q_2$	
$q_1$	$q_2$	$q_2$	$q_2$	$q_2$	$q_2$	$q_4$
$q_2$						$q_4$
$q_3$	$q_2$	$q_2$	$q_2$			$q_4$
$q_4$	$q_5$	$q_5$	$q_5$	$q_5$		
$q_5$	$q_6$	$q_6$	$q_6$	$q_6$	$q_6$	

Table 1: Ο πίνακας μεταβάσεων του DFA

Figure 1: DFA που αναγνωρίζει ενδείξεις ώρας σε 24ωρη μορφή, με αρχική κατάσταση  $q_0$  και τελική κατάσταση αποδοχής  $q_6$ .



Οι χαρακτήρες εισόδου έχουν χωριστεί σε ομάδες με στόχο τη μείωση του μεγέθους του πίνακα μεταβάσεων. Συγκεκριμένα:

- D\_UPTO1: Ψηφίο από 0 έως 1
- D\_2: Ψηφίο 2
- D\_3: Ψηφίο 3
- D\_UPTO5: Ψηφίο από 4 έως 5
- D\_ANY: Ψηφίο από 6 έως 9
- DELIMITER: Διαχωριστής (: ή .)

## 2 Προγραμματιστική Υλοποίηση DFA

Συνοπτικά, το αυτόματο υλοποιείται ως εξής:

1. Η συνάρτηση *scan*, με ορίσματα το κείμενο προς αναγνώριση, τον πίνακα μεταβάσεων και τον πίνακα καταστάσεων αποδοχής, σαρώνει το κείμενο, μέσω της *getchar*, όσο η τρέχουσα κατάσταση υπάρχει στον πίνακα μεταβάσεων. Μόλις φτάσει σε κατάσταση που υπάρχει στον πίνακα καταστάσεων αποδοχής, επιστρέφει το αντίστοιχο token, διαφορετικά `ERROR_TOKEN`.
2. Η συνάρτηση *getchar*, με ορίσματα το κείμενο προς αναγνώριση και έναν ακέραιο αριθμό, επιστρέφει την ομάδα χαρακτήρων στην οποία ανήκει ο χαρακτήρας στη θέση του κειμένου που προσδιορίζει ο ακέραιος. Αν ο χαρακτήρας δεν ανήκει σε κάποια ομάδα επιστρέφει `OTHER`, ενώ αν ο ακέραιος είναι εκτός ορίων του μήκους του κειμένου επιστρέφει `None`.
3. Οι πίνακες μεταβάσεων και καταστάσεων αποδοχής υλοποιούνται με dictionaries, όπως παρουσιάζονται παραπάνω. Το κείμενο προς αναγνώριση δίνεται από το χρήστη, περνιέται επαναληπτικά μαζί με τα dictionaries στη *scan* όσο αναγνωρίζονται χαρακτήρες, ώσπου να επιστραφεί `ERROR_TOKEN` ή να τελειώσει το κείμενο.

### 3 Αποτελέσματα εξόδου για έγκυρες και άκυρες μορφές ώρας

Figure 2: Αποτελέσματα εξόδου για *έγκυρες* μορφές ώρας.

```
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 23:45
token: TIME_TOKEN string: 23:45
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 7.22
token: TIME_TOKEN string: 7.22
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 09:08
token: TIME_TOKEN string: 09:08
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 3.00
token: TIME_TOKEN string: 3.00
```

Figure 3: Αποτελέσματα εξόδου για *άκυρες* μορφές ώρας.

```
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 17:4
unrecognized input at pos 5 of 17:4
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 25:12
unrecognized input at pos 2 of 25:12
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 9:66
unrecognized input at pos 3 of 9:66
alex@alex-VirtualBox:~/Desktop$ python3 scanner.py
give some input> 24:60
unrecognized input at pos 2 of 24:60
```

## References

- [1] N. Murugesan and O. Sundaram, “A general approach to dfa construction,” 06 2015.
- [2] E. Wallace, “Finite State Machine Designer,” 2010. [Online]. Available: <http://madebyevan.com/fsm/>. [Accessed: 19-Mar-2018].