

FSM

Εισαγωγή

Μια μηχανή πεπερασμένων καταστάσεων (finite state machines – FSM) είναι ένας μεταφραστής που δέχεται την περιγραφή της λειτουργίας ενός διαγράμματος μετάβασης σε μια συγκεκριμένη γλώσσα κατάλληλη για δήλωση μεταβάσεων, επιβεβαιώνει ότι το αυτόματο είναι πράγματι ντετερμινιστικό και προσομοιώνει τη λειτουργία του επάνω σε δεδομένα εισόδου που προσδιορίζεται από τον χρήστη, απαντώντας αν τα δεδομένα αναγνωρίζονται από το αυτόματο ή όχι.

Χρήση

Το εκτελέσιμο αρχείο λέγεται fsm.exe. Η περιγραφή του αυτόματου πρέπει να βρίσκεται σε ένα αρχείο με επέκταση .fsm . Για παράδειγμα, ας υποθέσουμε ότι το αρχείο test1.fsm περιέχει την περιγραφή ενός αυτόματου που αναγνωρίζει λέξεις από 0 και 1 με άρτιο αριθμό από 1. Ο πιο απλός τρόπος να το τρέξουμε είναι:

```
fsm.exe test1
```

Το **FSM** θα επιβεβαιώσει ότι το αυτόματο που περιγράφεται στο test1.fsm είναι συντακτικά σωστό και ντετερμινιστικό και μετά θα περιμένει να του δώσουμε μια λέξη για να μας πει αν είναι ή όχι αποδεκτή από τη γραμματική της γλώσσας που αναγνωρίζεται από το test1.fsm:

```
> fsm.exe test1
1111
YES
> fsm.exe test1
10
NO
```

Με την επιλογή **-list** το FSM τυπώνει τους κανόνες και με την επιλογή **-trace** τυπώνει λεπτομερώς τη λειτουργία αναγνώρισης:

```
> fsm.exe trace
test1
10
sz 1 -> so
so 0 -> so
so \n -> bad
NO
```

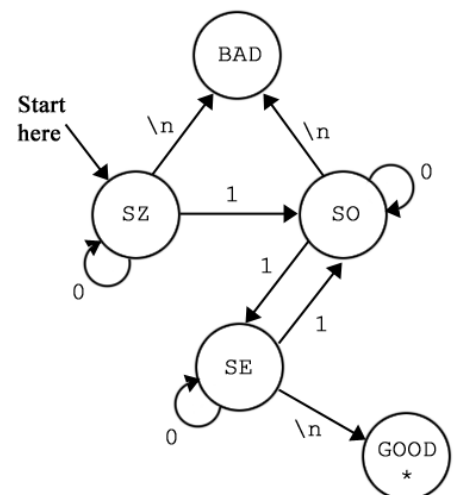
Προσοχή: Όταν πληκτρολογήσετε το input, πατήστε enter και μετά controlZ το οποίο στα Windows σημαίνει end-of-file.

Μόνο τότε παράγει output το FSM. Διαφορετικά, θεωρεί πως όποιος χαρακτήρας πληκτρολογείτε (ακόμη και το enter) είναι input.

Περιγραφή αυτόματου

Το δίπλα σχήμα δείχνει το διάγραμμα του αυτόματου που περιγράφεται στο test1.sfm.

Η αρχική κατάσταση είναι η sz. Από εκεί, αν διαβάσουμε ένα 1, πηγαίνουμε στην κατάσταση so (state odd – περιττός αριθμός από 1). Αν διαβάσουμε 0 μένουμε στην ίδια κατάσταση, ενώ με 1 μεταβαίνουμε στην κατάσταση se (state even - άρτιος αριθμός από 1). Ο μόνος τρόπος να γίνει αποδεκτό το input είναι αν είμαστε στην κατάσταση se και συναντήσουμε newline (ακολουθούμενο από end-of-file). Η περιγραφή είναι ως εξής:



```

START=SZ
SZ:  0 -> SZ
      1 -> SO
      \n -> BAD
SO:  0 -> SO
      1 -> SE
      \n -> BAD
SE:  0 -> SE
      1 -> SO
      \n -> GOOD
GOOD (OK) :

```

Γραμμή 1: Η γραμμή αυτή προσδιορίζει την αρχική κατάσταση. Μπορεί να παραληφθεί, στην οποία περίπτωση, αρχική κατάσταση θεωρείται αυτή που εμφανίζεται πρώτη στην επόμενη γραμμή.

Γραμμές 2-4: Οι μεταβάσεις από την κατάσταση sz με inputs 0, 1 και newline, ακριβώς όπως εμφανίζονται στο διάγραμμα.

Γραμμή 11: Η σύνταξη ΟΝΟΜΑ_ΚΑΤΑΣΤΑΣΗΣ(OK) προσδιορίζει μια κατάσταση εξόδου (accepting state).

Οι κανόνες διατύπωσης/περιγραφής του αυτομάτου είναι:

- Λευκές γραμμές επιτρέπονται και αγνοούνται
- Οι χαρακτήρες // εισάγουν σχόλια που λήγουν στο τέλος της γραμμής
- Τα ονόματα των καταστάσεων μπορούν να περιλαμβάνουν γράμματα, ψηφία, δολάρια. υπογραμμίσεις και τελείες. Πεζά και κεφαλαία γράμματα δεν ξεχωρίζονται.
- Η περιγραφή μπορεί να περιέχει μια γραμμή: START=state που προσδιορίζει την αρχική κατάσταση. Η γραμμή αυτή μπορεί να παραληφθεί, στην οποία περίπτωση, αρχική κατάσταση θεωρείται αυτή που εμφανίζεται πρώτη στην επόμενη γραμμή.
- Οι μεταβάσεις για κάθε κατάσταση ξεκινούν με το όνομα της κατάστασης ακολουθούμενο από ανω-κάτω τελεία. Οι μεταβάσεις μιας συγκεκριμένης κατάστασης δε χρειάζεται να βρίσκονται όλες μαζί (αλλά συνίσταται).
- Όταν εμφανίζεται η λέξη OK σε παρενθέσεις, ανάμεσα στο όνομα μιας κατάστασης και στην ανω-κάτω τελεία, τότε η κατάσταση είναι εξόδου (accepting). Μπορούν να υπάρχουν περισσότερες από μία καταστάσεις εξόδου.
- Μπορούν να υπάρχουν 0 ή περισσότερες μεταβάσεις για κάθε κατάσταση.
- Μια μετάβαση αποτελείται από τον χαρακτήρα (ή χαρακτήρες) που την προκαλεί, ένα βέλος -> και το όνομα της κατάστασης στην οποία μεταβαίνει.
- Ότ αν διάφοροι χαρακτήρες προκαλούν την ίδια μετάβαση, μπορούν να καταγραφούν μαζί, χωρισμένοι από κενό. Για παράδειγμα, η μετάβαση :

```
a b c -> S2
```

είναι ισοδύναμη με τις μεταβάσεις:

```

a -> S2
b -> S2
c -> S2

```

- Μια αλφαβητική σειρά χαρακτήρων μπορεί να προσδιοριστεί από τον πρώτο και τελευταίο χαρακτήρα με μια παύλα στη μέση. Για παράδειγμα, η παραπάνω μετάβαση θα μπορούσε να γραφτεί:

`ab-> S2`

Αυτό ισχύει και με ψηφία.

- Όταν χρησιμοποιείται ως χαρακτήρας, ο αστερίσκος `*` σημαίνει "οποιοσδήποτε χαρακτήρας δεν έχει ήδη προσδιοριστεί σε άλλη μετάβαση από αυτή την κατάσταση". Ως συνέπεια, δύο μεταβάσεις από την ίδια κατάσταση δε μπορούν να περιέχουν αστερίσκους (το πολύ μια μπορεί).
- Ειδικοί χαρακτήρες περιγράφονται ως εξής:
 - `\s` κενό
 - `\t` tab
 - `\n` newline
 - `*` αστερίσκος
 - `\\` κάθετος `\`
 - EOF τέλος αρχείου/input
- Η λέξη **none** μπορεί να εμφανιστεί σε μια μετάβαση αντί για τους χαρακτήρες μετάβασης. Σε αυτή την περίπτωση, η μετάβαση γίνεται αυτόματα χωρίς να καταναλωθεί input.
- Το αυτόματο πρέπει να είναι ντετερμινιστικό. __