Κεφάλαιο 1

Εισαγωγή και Μαθηματικό Υπόβαθρο

1.1 Αλφάβητα, λέξεις και γλώσσες

Ορισμός 1. (Αλφάβητο)
Αλφάβητο είναι ένα οποιοδήποτε μη κενό πεπερασμένο σύνολο, του οποίου τα στοιχεία ονομάζουμε *σύμβολα* ή *γράμματα* .

Παράδειγμα 1. Το *δυαδικό αλφάβητο* $\Sigma_2 = \{0, 1\}$. Υπό μία έννοια αυτό είναι το πλέον ενδιαφέρον αλφάβητο. Ένας λόγος είναι ότι αυτό είναι το αλφάβητο που χρησιμοποιείται στους σημερινούς υπολογιστές. Ο σημαντικότερος όμως λόγος είναι ότι το Σ_2 μπορεί να "μιμηθεί" αποτελεσματικά οποιοδήποτε άλλο αλφάβητο. Το τι σημαίνει αυτό ελπίζουμε να γίνει ξεκάθαρο αργότερα.

Παράδειγμα 2. Το αλφάβητο $\Sigma_{GR} = \{A, B, \Gamma, \ldots, \alpha, \beta, \gamma, \ldots, \omega, \ldots\}$ περιλαμβάνει όλα τα γράμματα και σημεία στίξης της Ελληνικής γλώσσας.

Ορισμός 2. (Λέξη)
Λέξη πάνω από ένα αλφάβητο Σ λέγεται μια πεπερασμένη ακολουθία

$$\alpha = \alpha_1 \ldots \alpha_n$$

γραμμάτων $\alpha_i \in \Sigma$, ενδεχομένως και κενή (δηλ. με $n = 0$).

Ο αριθμός $n \geq 0$ λέγεται *μήκος* της λέξης α και συμβολίζεται με $|\alpha|$.

Η κενή λέξη συμβολίζεται πάντα με το ϵ.

Με Σ^* συμβολίζουμε το σύνολο όλων των λέξεων πάνω από το αλφάβητο Σ.

Παράδειγμα 3. Η $w = 0101$ είναι μια λέξη πάνω από το δυαδικό αλφάβητο Σ_2, με μήκος $|w| = 4$.

Ορισμός 3. (Γλώσσα)
Μια *γλώσσα* L πάνω από ένα αλφάβητο Σ είναι ένα οποιοδήποτε σύνολο, πεπερασμένο ή άπειρο, λέξεων πάνω από το Σ. Ως συνήθως στα μαθηματικά με $|L|$ συμβολίζουμε τον πληθάριθμο της L.

Παράδειγμα 4. Το κενό σύνολο $\emptyset$ αποτελεί μια γλώσσα πάνω από οποιοδήποτε αλφάβητο Σ που θα το ονομάζουμε *κενή γλώσσα* .

Παράδειγμα 5. Αν είναι Σ ένα αλφάβητο τότε, καταχρηστικά, συμβολίζουμε πάλι με Σ τη γλώσσα πάνω από το Σ που αποτελείται από όλες τις δυνατές λέξεις με ακριβώς ένα γράμμα, το ίδιο δηλ. το αλφάβητο με μία έννοια.

Παράδειγμα 6. Το σύνολο Σ^* όλων των λέξεων πάνω από ένα αλφάβητο Σ είναι η *πλήρης* γλώσσα πάνω από το Σ. Για οποιοδήποτε Σ αυτή είναι μια άπειρη γλώσσα που περιέχει ως υπογλώσσες (υποσύνολα) όλες τις γλώσσες πάνω από το Σ.

Παράδειγμα 7. Το σύνολο E όλων των επωνύμων που εμφανίζονται τον τηλεφωνικό κατάλογο Κρήτης του 2003 αποτελεί μια γλώσσα πάνω από το αλφάβητο Σ_{GR} των γραμμάτων και σημείων στίξης της ελληνικής γλώσσας (υποθέτουμε εδώ ότι δεν χρησιμοποιούνται πουθενά γράμματα του λατινικού αλφαβήτου στον κατάλογο). Η E είναι μια πεπερασμένη αλλά μεγάλη γλώσσα.

Παράδειγμα 8. Το σύνολο

$$Q = \left\{ w \in \Sigma_2^* : \exists k \in \mathbf{N} : \ |w| = k^2 \right\}$$

είναι η γλώσσα όλων των δυαδικών λέξεων με μήκος που είναι τέλειο τετράγωνο. Προφανώς πρόκειται για μια άπειρη γλώσσα.

Παράδειγμα 9. Το σύνολο L_k όλων των λέξεων του Σ_2 με μήκος ακριβώς k, όπου k είναι ένας φυσικός αριθμός, είναι μια πεπερασμένη υπογλώσσα του Σ_2^*.

Άσκηση 1. Βρείτε το $|L_k|$. Αν M_k είναι η υπογλώσσα του Σ_2^* που αποτελείται από όλες τις λέξεις με μήκος το πολύ k, βρείτε το $|M_k|$.

Άσκηση 2. Πόσες λέξεις μήκους n υπάρχουν πάνω από ένα αλφάβητο με k γράμματα;

Παράδειγμα 10. Ας είναι Σ_A το αλφάβητο που αποτελείται από τα γράμματα του λατινικού αλφαβήτου, μικρά και κεφαλαία, τα δεκαδικά ψηφία, σημεία στίξης, παρενθέσεις, αγκύλες, τον *λευκό* (space) χαρακτήρα (να μη συγχέεται με την κενή λέξη ϵ), το χαρακτήρα αλλαγής γραμμής (carriage return), και γενικά όλα τα σύμβολα που εμφανίζονται σε ένα σύγχρονο πληκτρολόγιο υπολογιστή. Αυτό το αλφάβητο ονομάζεται και αλφάβητο ASCII[1].

Η γλώσσα L_C είναι εκείνο το σύνολο από λέξεις του Σ_A^* που αποτελούν συντακτικά σωστά προγράμματα στη γλώσσα προγραμματισμού C . Αν δεν είστε γνώστες της γλώσσας αυτής αλλά κάποιας άλλης, μπορείτε να ορίσετε την αντίστοιχη γλώσσα. Ένα πρόγραμμα λοιπόν δεν είναι τίποτε άλλο από μια λέξη σε ένα κατάλληλο αλφάβητο. Αν αυτό ξενίζει να θυμίσουμε ότι και οι χαρακτήρες αλλαγής γραμμής βρίσκονται μέσα στο αλφάβητο, και άρα μια λέξη του L_A μπορεί να περιέχει τα γράμματα ενός ολόκληρου αρχείου κειμένου.

Το L_C είναι μια άπειρη γλώσσα αφού δεν υπάρχει άνω όριο στο μέγεθος ενός συντακτικά σωστού προγράμματος, και άρα υπάρχουν άπειρα τέτοια.

Ορισμός 4. (Συγκόλληση λέξεων, πρόθεμα και επίθεμα λέξης)
Αν $\alpha = \alpha_1 \ldots \alpha_m$ και $\beta = \beta_1 \ldots \beta_n$ είναι δύο λέξεις πάνω από το αλφάβητο Σ με μήκη m και n, τότε ορίζουμε τη *συγκόλληση* (concatenation) $\alpha\beta$ να είναι η λέξη

$$\alpha\beta = \alpha_1 \ldots \alpha_m \beta_1 \ldots \beta_n,$$

που έχει μήκος το άθροισμα των μηκών.

Μια λέξη π λέγεται *πρόθεμα* (prefix) μιας λέξης α αν υπάρχει λέξη β τ.ώ. $\alpha = \pi\beta$. Ομοίως η π λέγεται *επίθεμα* (suffix) της α αν υπάρχει β τ.ώ. $\alpha = \beta\pi$.

Παρατήρηση 1. Προφανώς ισχύει πάντα $\alpha\epsilon = \epsilon\alpha = \alpha$, για κάθε λέξη α.

Είναι φανερό ότι μια λέξη με μήκος n έχει ακριβώς $n+1$ προθέματα και άλλα τόσα επιθέματα.

Παράδειγμα 11. Αν $\Sigma = \{a, b, c\}$ τότε η συγκόλληση των λέξεων abc και bb είναι η λέξη $abcbb$.

Παράδειγμα 12. Τα προθέματα της λέξης $abcd$ (δεν έχει σημασία σε ποιο αλφάβητο δουλεύουμε) είναι οι λέξεις $\epsilon, a, ab, abc, abcd$. Τα επιθέματα είναι οι λέξεις $abcd, bcd, cd, d, \epsilon$.

Παρατήρηση 2. Αν τ είναι γράμμα ενός αλφαβήτου Σ τότε συμβολίζουμε επίσης με τ τη γλώσσα πάνω από το Σ που περιέχει μόνο μια λέξη, τη λέξη τ, που έχει μόνο ένα γράμμα.

[1] Για την ακρίβεια ο *κώδικας* ASCII είναι ένα standard αντιστοίχισης όλων των συμβόλων που αναφέραμε παραπάνω στους αριθμούς 0-127. Το standard αυτό ακολουθείται σήμερα σε όλους τους υπολογιστές

Ορισμός 5. (Συγκόλληση γλωσσών)
Αν L_1, L_2 είναι δύο γλώσσες πάνω από το Σ ορίζουμε τη συγκόλληση L_1L_2 αυτών να είναι η γλώσσα

$$L_1L_2 = \{xy:\ x \in L_1, y \in L_2\}.$$

Αποτελείται δηλ. η L_1L_2 από όλες τις λέξεις που προκύπτουν ως συγκόλληση μιας λέξης από την L_1 με μια λέξη της L_2.

Ορίζουμε επίσης $L^0 = \{\epsilon\}$ και, για $n \geq 1$, $L^n = LL\cdots L$ (συγκόλληση της L με τον εαυτό της n φορές).

Τέλος ορίζουμε

$$L^* = \bigcup_{k=0}^{\infty} L^k$$

και

$$L^+ = \bigcup_{k=1}^{\infty} L^k.$$

Άσκηση 3. Περιγράψτε τη γλώσσα 1^* πάνω από το δυαδικό αλφάβητο Σ_2.

Άσκηση 4. Δείξτε ότι για κάθε γλώσσα L και φυσικούς αριθμούς m, n ισχύει $L^{m+n} = L^mL^n$.

Άσκηση 5. Αν L είναι μια οποιαδήποτε γλώσσα που δεν περιέχει την κενή λέξη ϵ, ποια είναι ακριβώς η διαφορά των γλωσσών L^* και L^+; Αλλάξει κάτι στην απάντηση αν $\epsilon \in L$;

Άσκηση 6. Βεβαιωθείτε ότι καταλαβαίνετε τι περιγράφει η γλώσσα

$$\{+, -, \epsilon\}1\{0,1\}^*.$$

(Πρόκειται για συγκόλληση τριών γλωσσών.) Όταν, όπως εδώ, δεν περιγράφουμε το αλφάβητο, αυτό συνάγεται από όλα τα σύμβολα που έχουν χρησιμοποιηθεί, στην προκειμένη περίπτωση δηλαδή $\Sigma = \{0, 1, +, -\}$.

1.2 Γραφήματα, δέντρα

Ορισμός 6. (Κατευθυνόμενο γράφημα)
Ένα *κατευθυνόμενο γράφημα* G είναι ένα ζεύγος από δύο σύνολα V και E και μια συνάρτηση (πολλαπλότητα ακμής) $m : E \to \mathbf{N}$. Το σύνολο V λέγεται σύνολο *κορυφών* ή *κόμβων* και το σύνολο *ακμών* E είναι ένα υποσύνολο του

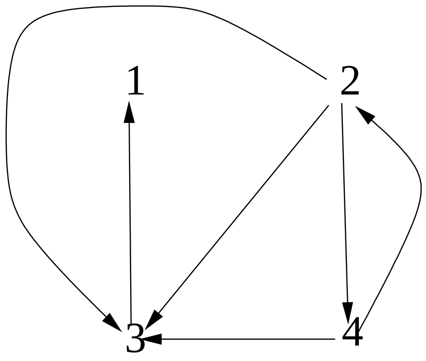
$$V \times V = \{(u, v):\ u, v \in V\}.$$

Αν το ζεύγος (u, v) ανήκει στο E λέμε ότι υπάρχει τουλάχιστον μια ακμή από το u στο v. Ένα κατευθυνόμενο γράφημα το ζωγραφίζουμε συνήθως σα μια συλλογή από κορυφές και ακμές που τις ενώνουν, όπως στο Σχήμα 1 παρακάτω. Λέγεται κατευθυνόμενο επειδή υπάρχει η έννοια της κατεύθυνσης πάνω στις ακμές. Οι δε ακμές μπορούν να ενώνουν και μια κορυφή με τον εαυτό της, όπως επίσης μπορούν να υπάρχουν και πολλαπλές ακμές που ενώνουν το ίδιο ζεύγος κορυφών. Το πόσες ακμές ενώνουν την κορυφή u με την κορυφή v (από το u προς το v) δίδεται από τον αριθμό $m(u, v)$, που είναι ένας θετικός ακέραιος.

Παράδειγμα 13. Στο Σχήμα 1 παριστάνεται το γράφημα $G = (V, E, m)$ με $V = \{1, 2, 3, 4\}$, $E = \{(2, 3), (2, 4), (3, 1), (4, 2), (4, 3$
και πολλαπλότητα m που είναι 1 για κάθε ακμή εκτός από τη (2,3) για την οποία έχουμε $m(2, 3) = 2$.

Πρέπει να τονίσουμε εδώ ότι σε ένα κατευθυνόμενο γράφημα (συνήθως θα λέμε απλώς γράφημα αντί για κατευθυνόμενο γράφημα όταν είναι ξακάθαρο τι εννούμε) δεν έχει καμία σημασία ο τρόπος που το σχεδιάζουμε πάνω στο χαρτί, αλλά αυτό που μας ενδιαφέρει είναι μόνο ποια κορυφή συνδέεται με ποια, και με τι πολλαπλότητα.

Ορισμός 7. (Διαδοχικές κορυφές)
Σε ένα κατευθυνόμενο γράφημα G η κορυφή v λέγεται *διαδοχική* της κορυφής u αν υπάρχει ακμή (u, v).



Σχήμα 1. Ενα κατευθυνόμενο γράφημα με 4 κορυφές και 5 ακμές. Μία από τις ακμές έχει πολλαπλότητα 2.

Ορισμός 8. (Μονοπάτια σε γραφήματα)
Σε ένα κατευθυνόμενο γράφημα *μονοπάτι* αποτελεί μια οποιαδήποτε πεπερασμένη ακολουθία κορυφών

$$\pi = u_0 u_1 \ldots u_n,$$

όπου η ακμή u_{i+1} είναι διαδοχική της u_i για κάθε $i = 0, \ldots, n-1$.

Ο αριθμός n λέγεται *μήκος* του μονοπατιού.

Αν οι κορυφές u_0 και u_n συμπίπτουν τότε το μονοπάτι λέγεται και *κύκλος*.

Παράδειγμα 14. Στο Σχήμα 1 το 24231 είναι ένα μονοπάτι μήκους 4. Το 424 είναι κύκλος.

Άσκηση 7. Έστω κατευθυνόμενο γράφημα G με n κορυφές και μονοπάτι στο G μέ μήκος $k \geq n$. Δείξτε ότι το μονοπάτι περιέχει ένα κύκλο.

Άσκηση 8. Αν σε ένα κατευθυνόμενο γράφημα G υπάρχει ένα μονοπάτι με μήκος μεγαλύτερο ή ίσο από το πλήθος των κορυφών τότε υπάρχουν άπειρα το πλήθος διαφορετικά μονοπάτια στο G.

Όταν δεν μας ενδιαφέρει η κατεύθυνση πάνω στις ακμές ενός γραφήματος και δεν επιτρέπουμε πολλαπλές ακμές από μια κορυφή σε μια άλλη τότε έχουμε μη κατευθυνόμενα, απλά γραφήματα. Στα γραφήματα αυτα γενικεύονται κατα προφανή τρόπο οι έννοιες των διαδοχικών κορυφών καθώς και των μονοπατιών.

Ορισμός 9. (Μη κατευθυνόμενο, απλό γράφημα)
Ένα *μη κατευθυνόμενο, απλό γράφημα* G είναι ένα ζεύγος από δύο σύνολα V και E. Το σύνολο V λέγεται σύνολο *κορυφών* ή *κόμβων* και το σύνολο *ακμών* E είναι ένα υποσύνολο του

$$\binom{V}{2} = \{\{u, v\} : \ u, v \in V, u \neq v\}.$$

Το E δηλ. απαρτίζεται από διμελή υποσύνολα του συνόλου των κορυφών. Αν το σύνολο $\{u, v\}$ ανήκει στο E τότε λέμε ότι η ακμή uv υπάρχει στο γράφημα G.

1.3 Σχέσεις πάνω σε σύνολα, σχέσεις ισοδυναμίας, μεταβατική κλειστότητα

Ορισμός 10. (Σχέση πάνω σε ένα σύνολο)
Σχέση R *πάνω σε ένα σύνολο* A είναι ένα οποιοδήποτε υποσύνολο $R \subseteq A \times A$.

Γράφουμε συνήθως $x\ R\ y$ αντί για $(x, y) \in R$.

Παράδειγμα 15. Ένα τετριμμένο παράδειγμα είναι η *πλήρης σχέση* $A \times A$, στην οποία δύο οποιαδήποτε στοιχεία του A σχετίζονται μεταξύ τους.

Παράδειγμα 16. Αν $A = \mathbf{R}$, το σύνολο των πραγματικών αριθμών, ορίζουμε τη σχέση $R = \{(x, y) \in \mathbf{R}^2 : x \leq y\}$. Αντί για $x\ R\ y$ γράφουμε κατά παράδοση $x \leq y$.

Παράδειγμα 17. Αν $A = \mathbf{Z}$, το σύνολο των ακεραίων, και $m \in \mathbf{Z}$ είναι ένας δοσμένος ακέραιος, $m \neq 0$, ορίζουμε $R = \{(x, y) \in \mathbf{Z}^2 : m|y - x\}$. Αντί για $x\ R\ y$ γράφουμε $x \sim_m y$ και διαβάζουμε "το x είναι ισουπόλοιπο mod m με το y". Με λίγη σκέψη βλέπουμε ότι αυτό συμβαίνει αν και μόνο αν τα x και y αφήνουν το ίδιο υπόλοιπο διαιρούμενα με m.

Παράδειγμα 18. Αν $A = \mathbf{Z}$, ορίζουμε $x\ R\ y$ αν και μόνο αν $2|x - y$ ή $3|x - y$.

Ορισμός 11. (Σχέσεις ανακλαστικές, συμμετρικές και μεταβατικές)
Μια σχέση R λέγεται *ανακλαστική* αν $x\ R\ x$ για κάθε $x \in A$, *συμμετρική* αν $x\ R\ y \Rightarrow y\ R\ x$ και, τέλος, *μεταβατική* αν $(x\ R\ y\ \&\ y\ R\ z) \Rightarrow x\ R\ z$.

Αν μια σχέση έχει και τις τρεις αυτές ιδιότητες τότε λέγεται σχέση *ισοδυναμίας*.

Παράδειγμα 19. Στο παράδειγμα 16 η σχέση είναι ανακλαστική και μεταβατική αλλά όχι συμμετρική. Στο παράδειγμα 17 η σχέση έχει και τις τρεις ιδιότητες και είναι συνεπώς μια σχέση ισοδυναμίας. Τέλος στο παράδειγμα 18 η σχέση είναι ανακλαστική και συμμετρική αλλά όχι και μεταβατική.

Άσκηση 9. Έστω G κατευθυνόμενο γράφημα με σύνολο κορυφών V. Ορίζουμε τη σχέση R στο V: $u\ R\ v$ αν υπάρχει μονοπάτι πάνω στο G που ξεκινά από το u και καταλήγει στο v. Είναι η σχέση R ανακλαστική, συμμετρική, μεταβατική·

Άσκηση 10. Γράφουμε $x\ R\ y$ για δυο ακεραίους x και y αν $x \mid y$ (ο x διαιρεί τον y). Δείξτε ότι η σχέση R είναι ανακλαστική και μεταβατική.

Άσκηση 11. Έστω R_1 και R_2 δύο σχέσεις ορισμένες πανω στο ίδιο σύνολο. Αν είναι ανακλαστικές (αντ. συμμετρικές, μεταβατικές) δείξτε ότι η σχεση $R_1 \cap R_2$ είναι επίσης ανακλαστική (αντ. συμμετρική, μεταβατική). Δείξτε το ίδιο και για την τομή

$$R = \bigcap_{j \in J} R_j$$

μιας οικογένειας σχέσεων, όπου το σύνολο δεικτών J δεν είναι κατ΄ ανάγκη πεπερασμένο, αλλά μπορεί να είναι οποιοδήποτε σύνολο.

Άσκηση 12. Ορίζουμε τη σχέση R στο σύνολο Σ_2^* (το σύνολο όλων των λέξεων που μπορεί κανείς να φτιάξει με τα γράμματα 0 και 1) ώστε να έχουμε xRy, για δύο λέξεις x και y, αν και μόνο αν η x είναι πρόθεμα της y. Είναι η σχέση R ανακλαστική, συμμετρική, μεταβατική;

Θεώρημα 1. (Ύπαρξη μεταβατικής κλειστότητας)
Έστω R μια σχέση πάνω σε ένα σύνολο A. Τότε υπάρχει μοναδική σχέση R^+ ορισμένη πάνω στο σύνολο A, που περιέχει την R, είναι μεταβατική και περιέχεται σε κάθε άλλη μεταβατική σχεση που περιέχει την R. Υπό αυτή την έννοια είναι η ελάχιστη μεταβατική σχέση που περιέχει την R *και την ονομάζουμε* μεταβατική κλειστότητα *ή* μεταβατική θήκη *της R.*

Απόδειξη. Ορίζουμε την κλάση $\mathcal{K}$ που απαρτίζεται από όλες τι σχέσεις X που περιέχουν την R και είναι ταυτόχρονα μεταβατικές. Η κλάση $\mathcal{K}$ είναι σίγουρα μη κενή μια και περιέχει την πλήρη σχέση $A \times A$. Ορίζουμε τώρα τη σχέση R^+ ως

$$R^+ = \bigcap_{X \in \mathcal{K}} X.$$

Είναι προφανές ότι $R \subseteq R^+$, και ότι καθε μεταβατική σχέση που περιέχει την R, δηλ. κάθε σχέση που ανήκει στην κλάση $\mathcal{K}$, περιέχει την R^+. Το ότι η R^+ είναι η ίδια μεταβατική προκύπτει από την Άσκηση 11.
□

Παρατήρηση 3. Υπάρχει και ένας άλλος τρόπος να δούμε τη σχέση R^+, εξ ίσου σημαντικός με τον τυπικό ορισμό. Το να μην είναι μια σχέση R μεταβατική σημαίνει ότι υπάρχουν ζεύγη $(a, b), (b, c) \in R$ αλλά $(a, c) \notin R$. Για να μετατρέψουμε λοιπόν τη σχέση R στη σχέση R^+, την 'ελάχιστη μεταβατική σχέση που περιέχει την R', προσθέτουμε στη σχέση R τα ζευγάρια αυτά που λείπουν, μέχρι να μη χρειάζεται να προσθέσουμε άλλα. Φυσικά, αυτή η διαδικασία που μόλις περιγράψαμε, για να έχει νόημα, πρέπει κάποια στιγμή να τελειώσει. Και αυτό δεν είναι πρόβλημα για σχέσεις R που είναι πεπερασμένες (τέτοιες είναι, για παράδειγμα, όλες οι σχέσεις ορισμένες σε πεπερασμένο σύνολο A) αλλά όταν θέλουμε να ορίσουμε το R^+ για άπειρες σχέσεις R πρέπει να χρησιμοποιήσουμε τον ορισμό που δώσαμε πιο πάνω με την, εν δυνάμει, άπειρη τομή.

Άσκηση 13. Αν μια σχέση R είναι ανακλαστική και συμμετρική τότε η R^+ είναι σχέση ισοδυναμίας.

Παράδειγμα 20. Έστω A το σύνολο όλων των ανθρώπων που έχουν ζήσει ποτέ. Έστω R η σχέση της γέννησης. Έχουμε δηλ. $x\ R\ y$ για δύο ανθρώπους x και y αν ο x είναι γονέας του y. Είναι φανερό ότι η σχέση R δεν είναι μεταβατική (αφού, π.χ. οι παππούδες δεν έχουν γεννήσει τα εγγόνια). Η μεταβατική κλειστότητα R^+ της R είναι η σχέση των απογόνων. Έχουμε δηλ. $x\ R^+\ y$ αν και μόνο αν ο y είναι απόγονος του x (ο y έχει γεννηθεί από κάποιον που έχει γεννηθεί από κάποιον, κλπ, που έχει γεννηθεί από τον x).

Παράδειγμα 21. Δείχνουμε τώρα ότι η μεταβατική κλειστότητα της σχέσης 18, παραπάνω, είναι η *πλήρης* σχέση στους ακεραίους, δηλ. κάθε δύο ακέραιοι σχετίζονται μεταξύ τους. Πράγματι, αν πούμε R τη σχέση 18, τότε $x\ R^+\ x$ για κάθε $x \in \mathbf{Z}$ αφού το 0 διαιρείται από το 2 (και το 3). Επίσης αποδεικνύουμε με επαγωγή, ως προς το θετικό αριθμό k, ότι για κάθε $x \in \mathbf{Z}$ ισχύει $x\ R^+\ x+k$ και $x+k\ R^+\ x$. Δείχνουμε μόνο το πρώτο μια και το δεύτερο είναι εντελώς αντίστοιχο. Για $k = 1$ πρέπει να δείξουμε ότι $x\ R^+\ x+1$. Αλλά $x\ R\ x+3$ (διαφορά ίση με 3) και $x+3\ R\ x+1$ (διαφορά ίση με 2) και επειδή $R \subseteq R^+$ έχουμε επίσης $x\ R^+\ x+3$, $x+3\ R^+\ x+1$ και από τη μεταβατικότητα της R^+ έχουμε $x\ R^+\ x+1$.

Υποθέτουμε τώρα $x\ R^+\ x+k$ για κάποιο k και για κάθε x και πάμε να δείξουμε $x\ R^+\ x+k+1$. Αλλά έχουμε $x\ R^+\ x+k$ και από το βήμα 1 έχουμε $x+k-1\ R^+\ x+k$. Από μεταβατικότητα έχουμε $x\ R^+\ x+k+1$, που τελειώνει την απόδειξη.

Άσκηση 14. Για δυο ακεραίους x και y γράφουμε $x\ R\ y$ αν το $|x-y|$ ισούται με 15 ή 13. Ποια η μεταβατική κλειστότητα της σχέσης R;

Άσκηση 15. Έστω $\Sigma = \{a, b, c\}$ και R η σχέση πάνω στο Σ^* που ορίζεται ως $x\ R\ y$ αν και μόνο αν $y = xa$ ή $y = xb$. Περιγράψτε τη μεταβατική κλειστότητα της R.

Ορισμός 12. (Ισοδύναμα στοιχεία, κλάσεις ισοδυναμίας)
Αν R είναι μια σχέση ισοδυναμίας πάνω σε ένα σύνολο A, $x, y \in A$, και $x\ R\ y$ τότε τα x και y ονομάζονται R-ισοδύναμα στοιχεία ή απλώς ισοδύναμα.

Ένα μη κενό υποσύνολο $K \subseteq A$ ονομάζεται κλάση ισοδυναμίας της σχέσης R αν κάθε δύο στοιχεία του K είναι μεταξύ τους R-ισοδύναμα και, επί πλέον, αν $x \in K$ και $x\ R\ y$ τότε έπεται ότι και $y \in K$.

Άσκηση 16. Ποιες από τις παρακάτω σχέσεις στο σύνολο $\{1, 2, 3, 4\}$ είναι σχέσεις ισοδυναμίας;

1. $\{(1, 1), (2, 2), (3, 3), (4, 4)\}$.

2. $\{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$.

3. $\{(1, 1), (1, 2), (1, 4), (2, 2), (2, 4), (3, 3), (4, 1), (4, 2), (4, 4)\}$.

Θεώρημα 2. (Διαμέριση του συνόλου σε κλάσεις ισοδυναμίας)
(α) Αν R είναι μια σχέση ισοδυναμίας πάνω σε ένα σύνολο A και K_1, K_2 είναι δύο διαφορετικές κλάσεις ισοδυναμίας της R τότε $K_1 \cap K_2 = \emptyset$.
(β) Το σύνολο A γράφεται ως ξένη ένωση κλάσεων ισοδυναμίας. Έτσι δύο στοιχεία του A είναι μεταξύ τους ισοδύναμα αν και μόνο αν ανήκουν στην ίδια κλάση ισοδυναμίας.

Απόδειξη. (α) Έστω $x \in K_1 \cap K_2$. Αν $K_1 \neq K_2$ τότε υπάρχει κάποιο στοιχείο του ενός συνόλου που δεν ανήκει στο άλλο. Χωρίς βλάβη της γενικότητας ας υποθέσουμε πως $y \in K_1 \setminus K_2$. Όμως επειδή $x, y \in K_1$ και

K_1 είναι κλάση έπεται ότι $x \; R \; y$. Αλλά αυτό συνεπάγεται ότι και $y \in K_2$ μια και η K_2 είναι κλάση, άτοπο γιατί υποθέσαμε αλλιώς. Άρα τα K_1, K_2 είναι μεταξύ τους ξένα.

(β) Έστω X το σύνολο όλων των κλάσεων ισοδυναμία της R. Ισχύει τότε

$$A = \bigcup_{K \in X} K,$$

μια και κάθε στοιχείο του A ανήκει σε κάποια κλάση ισοδυναμίας (για παράδειγμα στο σύνολο όλων των στοιχείων του A που είναι ισοδύναμα προς αυτό – αυτό είναι μια κλάση ισοδυναμίας). Τα σύνολα που συμμετέχουν στην παραπάνω ένωση είναι ανά δύο ξένα, σύμφωνα με το (α), άρα το A διαμερίζεται στις κλάσεις ισοδυναμίας του ως προς τη σχέση ισοδυναμίας R.
$\square$

Ορισμός 13. (Δείκτης μια σχέσης ισοδυναμίας)
Αν R είναι μια σχέση ισοδυναμίας πάνω σε ένα σύνολο A ο πληθάριθμος του συνόλου των κλάσεων ισοδυναμίας ονομάζεται *δείκτης* της σχέσης R, και μπορεί φυσικά να είναι πεπερασμένος ή άπειρος.

Άσκηση 17. Σε ένα δωμάτιο βρίσκονται άνδρες και γυναίκες και φοράνε μπλούζες χρώματος κόκκινου, πράσινου ή μπλέ. Αν x και y είναι δύο άτομα γράφουμε $x \; R \; y$ για να δηλώσουμε ότι τα δύο άτομα είναι του ίδιου φύλου και φοράνε το ίδιο χρώμα. Δείξτε ότι η R είναι μια σχέση ισοδυναμίας και περιγράψτε τις κλάσεις ισοδυναμίας. Ποιες είναι οι κλάσεις ισοδυναμίες και πόσες το πολύ είναι;

Άσκηση 18. Σε ένα κράτος της Πολυνησίας που απαρτίζεται από 14 νησιά δυο σημεία του εδάφους θεωρούνται ότι σχετίζονται από τη σχέση R αν μπορεί κανείς να πάει από το ένα στο άλλο περπατώντας (δεν υπάρχουν ανωμαλίες στο έδαφος). Δείξτε ότι η R είναι σχέση ισοδυναμίας και βρείτε το δείκτη της.

Άσκηση 19. Να βρεθεί ο δείκτης της σχέσης 17 πάνω στους ακεραίους και να περιγραφούν οι κλάσεις ισοδυναμίας.

Άσκηση 20. Έστω f μια συνάρτηση με πεδίο ορισμού το $\mathbf{R}$ και πεδίο τιμών το σύνολο $\{1, \ldots, 10\}$. Για $x, y \in \mathbf{R}$ ορίζουμε $x \; R \; y$ αν και μόνο αν $f(x) = f(y)$. Δείξτε ότι η R είναι σχέση ισοδυναμίας, προσδιορίστε τις κλάσεις, κα πείτε ποια η σχέση του δείκτη με το σύνολο τιμών $f(\mathbf{R})$.

Άσκηση 21. Έστω $G = (V, E)$ ένα μη κατευθυνόμενο, απλό γράφημα. Για $u, v \in V$ γράφουμε $u \; R \; v$ αν και μόνο αν $u = v$ ή $u \neq v$ και $\{u, v\} \in E$. Περιγράψτε τις κλάσεις ισοδυναμίας της σχέσης R^+, αφού εξηγήσετε γιατί αυτή είναι σχέση ισοδυναμίας.

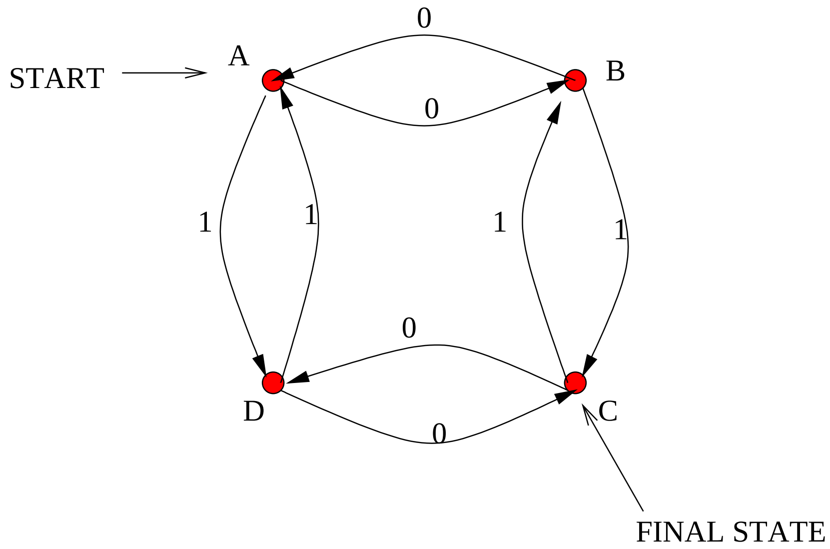
Κεφάλαιο 2

Αυτόματα

2.1 Ντετερμινιστικά Αυτόματα

Ένα Ντετερμινιστικό Αυτόματο (Deterministic Finite Automaton ή DFA) είναι ουσιαστικά ένα κατευθυνόμενο γράφημα, του οποίου οι κορυφές Q ονομάζονται *καταστάσεις* (states) και από κάθε κορυφή φεύγει ακριβώς μια ακμή για κάθε γράμμα του αλφαβήτου Σ. Υπάρχει μια διακεκριμένη κατάσταση q_0, η *αρχική κατάσταση* και ένα μη-κενό σύνολο F από *τελικές καταστάσεις*.

Δείτε π.χ. ένα τέτοιο αυτόματο στο Σχήμα 2.



Σχήμα 2. Ένα απλό ντετερμινιστικό αυτόματο

Ορισμός 14. (Ντετερμινιστικό Αυτόματο)
Ένα *ντετερμινιστικό αυτόματο* είναι μια πεντάδα

$$(Q, \Sigma, \delta, q_0, F)$$

όπου

- Q είναι ένα πεπερασμένο σύνολο *καταστάσεων*,

- Σ είναι ένα πεπερασμένο αλφάβητο,
- δ είναι η *συνάρτηση μετάβασης* (transition function) με πεδίο ορισμού το $Q \times \Sigma$ και πεδίο τιμών το Q,
- $q_0 \in Q$ είναι μια από τις καταστάσεις που ονομάζεται *αρχική*, και
- $F \subseteq Q$ είναι το σύνολο των *τελικών* καταστάσεων.

Παράδειγμα 22. Στο αυτόματο που φαίνεται στο Σχήμα 2 έχουμε $Q = \{A, B, C, D\}$, $\Sigma = \{0, 1\}$, $q_0 = A$, $F = \{C\}$.

Για να μιλήσουμε για τη συνάρτηση μετάβασης δ θα πρέπει πρώτα να αναφερθούμε στον τρόπο "λειτουργίας" του αυτομάτου. Ένα αυτόματο χρησιμεύει για να *αναγνωρίζει*, όπως λέμε, μια γλώσσα $L \subseteq \Sigma^*$. Η διαδικασία αναγνώρισης είναι η εξής.

Έστω λέξη $w \in \Sigma^*$, $w = w_1 \ldots w_n$, μήκους n ($w_i \in \Sigma$).

- Το αυτόματο αρχίζει τη λειτουργία του στην αρχική κατάσταση q_0.
- Διαβάζει έπειτα τα γράμματα της λέξης ένα προς ένα, από τα αριστερά προς τα δεξιά πάντα. Πρώτο διαβάζεται το γράμμα w_1 και τελευταίο το w_n.
- Μόλις διαβάσει το γράμμα $\alpha \in \Sigma$ και ευρισκόμενο στην κατάσταση q το αυτόματο *μεταβαίνει* στην κατάσταση r αν και μόνο αν η ακμή $q \to r$ έχει ετικέτα (label) ίση με α. Για κάθε κατάσταση του αυτομάτου και κάθε γράμμα του αλφαβήτου υπάρχει εξ ορισμού ακριβώς μια ακμή που ξεκινά από την κατάσταση αυτή και έχει ετικέτα αυτό το γράμμα. Ισχύει τότε για τη συνάρτηση μετάβασης $\delta(q, \alpha) = r$, χρησιμεύει δηλ. η συνάρτηση μετάβασης για να μας προσδιορίσει σε ποια κατάσταση πάει το αυτόματο αν βρίσκεται σε μια δεδομένη κατάσταση q και διαβάσει ένα συγκεκριμένο γράμμα α. Όλη η συνδεσμολογία του αυτομάτου δηλ. είναι κωδικοποιημένη στη συνάρτηση δ.
- Αφού διαβάσει το αυτόματο το τελευταίο γράμμα της λέξης και κάνει την τελευταία του μετάβαση δηλώνει ότι αποδέχεται τη λέξη αν και μόνο αν βρίσκεται σε τελική κατάσταση, σε μια κατάσταση δηλ. του συνόλου F. Αλλιώς η λέξη απορρίπτεται.

Ορισμός 15. (Γλώσσα ενός DFA)
Το σύνολο των λέξεων που αποδέχεται το αυτόματο M ονομάζεται *η γλώσσα που αναγνωρίζει το αυτόματο* και συμβολίζεται με $L(M)$.

Παράδειγμα 23. Ποια είναι η γλώσσα L που αναγνωρίζεται από το αυτόματο της εικόνας 2;

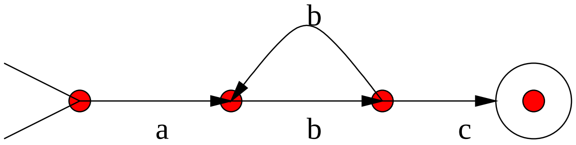
Δε θέλει και πολλή σκέψη για να πειστούμε ότι στην L ανήκουν ακριβώς εκείνες οι λέξεις του $\{0, 1\}^*$ που έχουν περιττό αριθμό από 0 και περιττό αριθμό από 1. Για να το δούμε αυτό παρατηρούμε ότι οποτεδήποτε το αυτόματο βρίσκεται σε μια από τις δύο αριστερές καταστάσεις το πλήθος των μηδενικών που έχει διαβάσει είναι άρτιο. Αυτό γίνεται γιατί αυτή η πρόταση ισχύει προφανέστατα τη χρονική στιγμή 0, αφού τότε δεν έχει διαβάσει κανένα μηδενικό και βρίσκεται στην αρχική κατάσταση A που είναι στην αριστερή μεριά. Οποτεδήποτε διαβάσει επίσης κάποιο μηδενικό το αυτόματο αλλάζει μεριά και διατηρείται έτσι η ιδιότητα αυτή. Ομοίως επιχειρηματολογώντας βλέπουμε ότι το αυτόματο βρίσκεται σε μια από τις δύο πάνω καταστάσεις (A και B) αν και μόνο αν έχει διαβάσει άρτιο αριθμό από 1. Έτσι, το να είναι το αυτόματο στην κατάσταση C (τελική) σημαίνει ότι έχει δει περιττό αριθμό από ένα και περιττό από μηδέν, αφού η C είναι κάτω (περιττοί άσσοι) και δεξιά (περιττά μηδενικά).

Αν π.χ. θελήσουμε να έχουμε ένα αυτόματο που θα αναγνωρίζει ακριβώς εκείνες τις λέξεις του $\{0, 1\}^*$ που έχουν *περιττά μηδενικά ή άρτιους άσσους* η μόνη αλλαγή που χρειάζεται να κάνουμε στην περιγραφή του αυτομάτου είναι να αλλάξουμε το σύνολο F των τελικών καταστάσεων και να το θέσουμε ίσο με το $\{A, B, C\}$.

Άσκηση 22. Σχεδιάστε ένα DFA που να αναγνωρίζει εκείνες τις λέξεις πάνω από το $\Sigma = \{0, 1\}$ που έχουν άρτιο πλήθος άσσων και πλήθος μηδενικών που είναι πολλαπλάσιο του 3.

Άσκηση 23. Σχεδιάστε ένα DFA που αναγνωρίζει εκείνες τις λέξεις πάνω από το $\Sigma = \{0, 1\}$ που αρχίζουν από 11011.

Παράδειγμα 24. Στο σχήμα 3 βλέπουμε ένα αυτόματο που αναγνωρζει τη γλώσσα $a\{bb\}^*bc$, δηλ. όλες εκείνες τις λέξεις στο αλφάβητο $\Sigma = \{a, b\}$ που αρχίζουν με a, ακολουθεί ένας οποιοσδήποτε αριθμός (ακόμη και μηδέν) από αντίγραφα της λέξης bb, και τελειώνουν με τη λέξη bc.



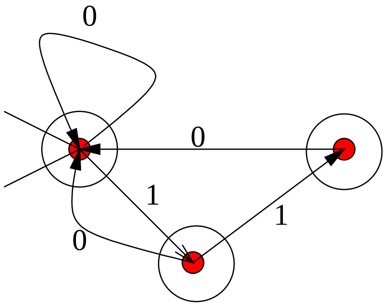
Σχήμα 3. DFA για τη γλώσσα $a\{bb\}^*bc$

Συμβολισμός: Στο Σχήμα 3, όπως και παρακάτω, συμβολίζουμε την αρχική κατάσταση με μια ανοιχτή γωνία (σα "χωνί", απ΄ όπου μπαίνουμε στο αυτόματο) και κάθε τελική κατάσταση μπαίνει μέσα σ'ενα κύκλο.

Συμβολισμός: Προσέξτε ότι στο αυτόματο του Σχήματος 3 δεν ισχύει η αρχική μας απαίτηση ότι από κάθε κορυφή πρέπει να φεύγει ακριβώς μια ακμή για κάθε γράμμα του αλφαβήτου (ώστε ό,τι και να διαβάσουμε όταν είμαστε σε κάποια κατάσταση να έχουμε που να πάμε). Για παράδειγμα, για τη δεύτερη κορυφή από αριστερά δεν υπάρχει ακμή με ετικέτα a που να ξεκινάει από αυτή. Η σύμβαση που ακολουθούμε είναι ότι αν διαβάσουμε ένα σύμβολο και βρισκόμαστε σε μια κατάσταση από την οποία δεν ξεκινάει ακμή με ετικέτα αυτό το σύμβολο, τότε η λέξη *δεν αναγνωρίζεται* από το αυτόματο.

Η σύμβαση αυτή ακολουθείται καθαρά για λόγους αισθητικής και κατανόησης του σχεδίου και δεν αλλάζει απολύτως τίποτα στην ουσία. Κάθε αυτόματο που ακολουθεί τη σύμβαση αυτή μπορεί εύκολα να μετατραπεί σε ένα αυτόματο που δε χρησιμοποιεί αυτή τη σύμβαση και έχει πράγματι μια ακμή για κάθε γράμμα από κάθε κορυφή. Απλά δηλώνουμε μια νέα *μη τελική* κατάσταση K, την κατάσταση καταστροφής, όπως επιλέγουμε να την ονομάσουμε, και από κάθε άλλη κατάσταση του αυτομάτου που δεν έχει ακμή από αυτή με ετικέτα έστω x ορίζουμε μια τέτοια ακμή προς την K. Όλες οι ακμές από το K επιστρέφουν πάλι στο K. Είναι εύκολο να δούμε ότι οι ίδιες ακριβώς λέξεις αναγνωρίζονται από το αρχικό και το νέο αυτόματο.

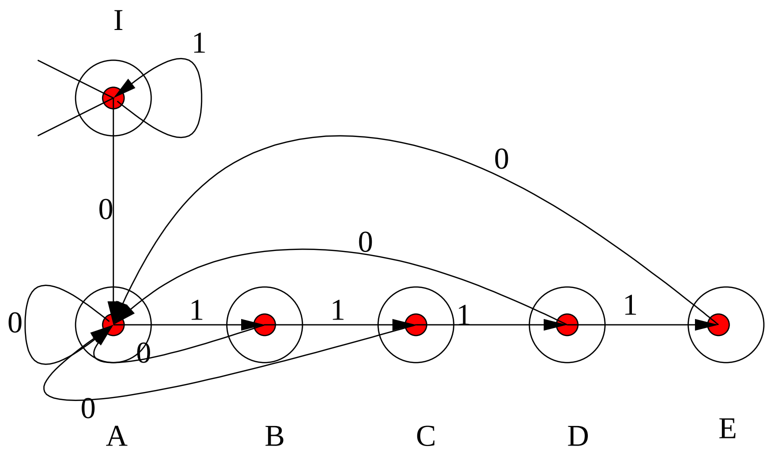
Παράδειγμα 25. Ακολουθεί στο Σχήμα 4 ένα αυτόματο που αναγνωρίζει τη γλώσσα L εκείνων των λέξεων του $\{0, 1\}^*$ που έχουν μέσα το πολύ δύο διαδοχικά 1. Ισοδύναμα, είναι εκείνες οι λέξεις στις οποίες *δεν εμφανίζεται* η μορφή 111.



Σχήμα 4. DFA για τις λέξεις χωρίς τρία διαδοχικά 1

Στο αυτόματο αυτό όλες οι καταστάσεις είναι τελικές, αλλά αυτό δεν αποτελεί αντίφαση, αφού ακολουθούμε τη σύμβαση της άνω παρατήρησης. Έτσι υπάρχουν πράγματι λέξεις που δεν αναγνωρίζονται από το αυτόματο, μια και κάποιες κορυφές δεν έχουν ακμές που ξεκινάνε από αυτές και για το 0 και το 1. Η μόνη τέτοια κορυφή είναι η δεξιά, και, όντως, η μόνη περίπτωση να απορριφθεί μια λέξη από το αυτόματο είναι να είμαστε στη δεξιά κατάσταση και να διαβάσουμε 1, πράγμα το οποίο συμβαίνει αν και μόνο αν υπάρχουν τρία διαδοχικά 1 στη λέξη.

Παράδειγμα 26. Το παράδειγμα αυτό είναι κάπως πιο ενδιαφέρον: L είναι η γλώσσα εκείνων των λέξεων του $\{0,1\}^*$ που είναι τέτοιες ώστε, *μετά το πρώτο 0* υπάρχει τουλάχιστον ένα 0 σε κάθε πεντάδα διαδοχικών γραμμάτων της λέξης. Για παράδειγμα, οι λέξεις 11111111111, 11111110111101111 είναι στην L ενώ η λέξη 111111011111 δεν είναι (η τελευταία πεντάδα δεν έχει 0). Το αυτόματο δίνεται στο Σχήμα 5. Πώς πρέπει να



Σχήμα 5. DFA για τις λέξεις με τουλάχιστον ένα 0 σε κάθε πεντάδα, μετά το πρώτο 0

σκεφθεί κανείς για να καταλάβει πώς δουλεύει το συγκεκριμένο αυτόματο;

Το κλειδί σε αυτή την περίπτωση, όπως και στις περισσότερες, είναι να αποδώσουμε κάποιο νόημα στην πρόταση 'βρισκόμαστε στην τάδε κατάσταση'. Για παράδειγμα, στην κατάσταση I βρισκόμαστε ακριβώς όταν δεν έχουμε διαβάσει ακόμη το πρώτο 0 της λέξης. Η πρόταση αυτή ισχύει κατ΄ αρχήν (πριν έχουμε διαβάσει τίποτε) και διατηρείται από οποιαδήποτε μετάβαση, αφού φεύγουμε από την κατάσταση I ακριβώς μόλις διαβάσουμε το πρώτο μηδενικό και δεν ξαναγυρνάμε σε αυτή. Είναι τώρα φανερό, έχοντας ξεκαθαρίσει τι σημαίνει να βρισκόμαστε στην κατάσταση I, ότι αυτή πρέπει να είναι τελική κατάσταση, μια και αν ποτέ δε διαβάσουμε κάποιο 0, ισχύει ο κανόνας που βάλαμε στην αρχή για το ποιες λέξεις ανήκουν στην L και άρα πρέπει να δεχτούμε τη λέξη.

Κάνουμε τώρα την εξής παρατήρηση. Ας πούμε πως ένας άνθρωπος, με λίγη μνήμη, έχει επιφορτισθεί με το καθήκον να αναγνωρίζει ακριβώς αυτές τις λέξεις. Αυτός μαθαίνει τα σύμβολα της λέξης ένα-ένα, πρώτα από τα αριστερά, όπως ακριβώς και το αυτόματο, και πρέπει κάποια στιγμή να πεί αν αυτή η λέξη ανήκει ή όχι στην L. Επειδή ακριβώς ο άνθρωπος αυτός έχει λίγη μνήμη, κι επειδή οι λέξεις τις οποίες κρίνει μπορούν να έχουν οσοδήποτε μεγάλο μήκος, αποφασίζει αντί ανά πάσα στιγμή να θυμάται ολόκληρη τη λέξη που έχει δει μέχρι τότε, να θυμάται απλώς *πόσο χρόνο πριν είδε το τελευταίο μηδενικό*. Είναι φανερό ότι αυτή η πληροφορία του αρκεί για να αποφασίσει. Θα απορρίψει δε τη λέξη αν και μόνο αν αυτός ο χρόνος ξεπεράσει το 5 σε κάποια χρονική στιγμή, γιατί ακριβώς τότε έχει δει μια πεντάδα γεμάτη με 1.

Μετά από αυτή την παρατήρηση το νόημα της κατάστασης A είναι ότι βρίσκεται εκεί αν έχει δεί 0 ακριβώς πριν από χρόνο 1. Το νόημα της κατάστασης B ομοίως είναι ότι έχει δεί το τελευταίο 0 ακριβώς πριν από χρόνο 2, κ.ο.κ., ενώ στην κατάσταση E βρισκόμαστε αν έχουμε δεί το τελευταίο 0 ακριβώς 5 χρονικές στιγμές πριν. Είναι φανερό τώρα γιατί η κατάσταση E δεν έχει μετάβαση που ξεκινά από αυτή με 1: αν είδαμε το τελευταίο 0 χρόνο 5 πίσω και μας έρθει 1 πάει να πει πως έχουμε μόλις συμπληρώσει μια πεντάδα γεμάτη με 1. Άρα δεν υπάρχει λόγος να συνεχίσουμε γιατί, ό,τι και να δούμε από δω και πέρα, η λέξη πρέπει να απορριφθεί.

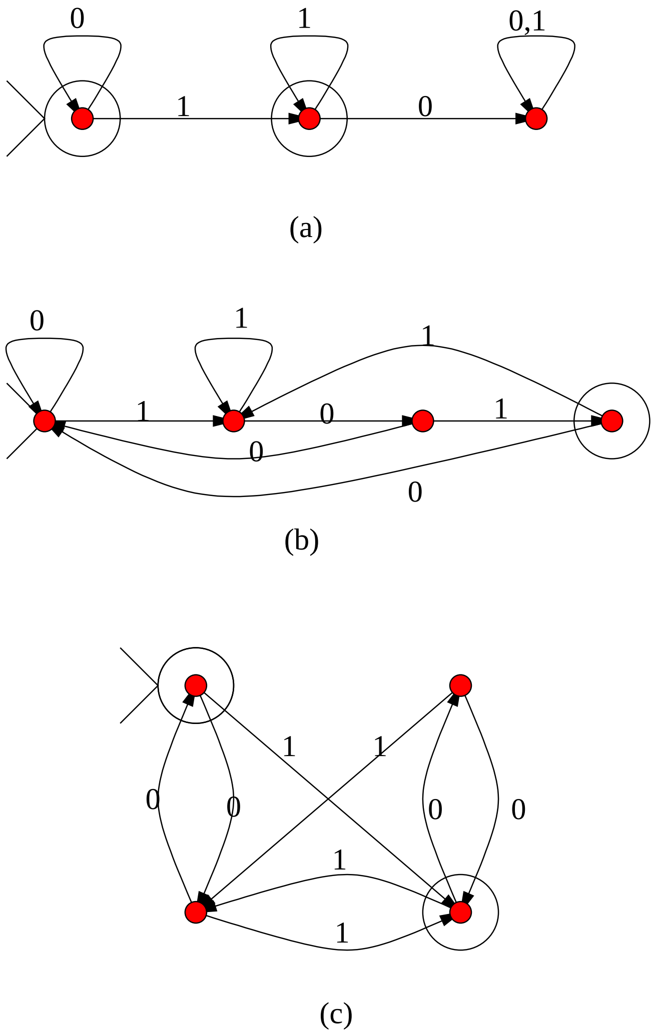
Έυκολα μπορούμε τώρα να ελέγξουμε ότι οι διάφορες μεταβάσεις στο σχήμα 5 είναι φτιαγμένες ακριβώς ώστε ακολουθώντας αυτές να διατηρείται το νόημα της κάθε κατάστασης, και άρα το αυτόματο δουλεύει. Αν, για παράδειγμα, βρισκόμαστε σε μια οποιαδήποτε από τις καταστάσεις $A,\ldots,E$ και διαβάσουμε 0 πρέπει να μεταβούμε στην κατάσταση A αφού μετά τη μετάβασή μας θα έχουμε διαβάσει 0 ακριβώς χρόνο 1 πριν, και αυτό είναι το νόημα της κατάστασης A.

Άσκηση 24. Κατασκευάστε ένα DFA που αναγνωρίζει εκείνες τις λέξεις πάνω από το $\Sigma = \{0,1\}$ που έχουν μήκος τουλάχιστον 5 και το 5ο γράμμα τους από τα αριστερά είναι 1.

Άσκηση 25. Κατασκευάστε ένα DFA που αναγνωρίζει εκείνες τις λέξεις πάνω από το $\Sigma = \{0,1\}$ που έχουν μήκος τουλάχιστον 5 και το 5ο γράμμα τους από τα δεξιά είναι 1.

Άσκηση 26. L είναι η γλώσσα εκείνων των λέξεων του $\{0,1\}^*$ που είναι τέτοιες ώστε, *μετά το πρώτο 0* υπάρχουν τουλάχιστον δύο 0 σε κάθε πεντάδα διαδοχικών γραμμάτων της λέξης. Κατασκευάστε ένα DFA που αναγνωρίζει την L.

Άσκηση 27. Περιγράψτε με όσο το δυνατό πιο απλό τρόπο, χωρίς ιδιαίτερους συμβολισμούς, ποιες γλώσσες είναι που αναγνωρίζονται από τα αυτόματα του Σχήματος 6.



Σχήμα 6. Ποιες γλώσσες αναγνωρίζουν τα αυτόματα αυτά;

Άσκηση 28. Κατασκευάστε DFA για τις γλώσσες του $\{0,1\}^*$:

1. Λέξεις που τελειώνουν σε 00
2. Λέξεις που περιέχουν τρία διαδοχικά 0

Άσκηση 29. Κατασκευάστε DFA για τις γλώσσες

(α) 00^*1^*1, και

(β) $\{w \in \{a,b\}^* : \quad 6 \mid A(w) + 2B(w)\}$,

όπου $A(w)$, και $B(w)$ είναι το πλήθος των a και των b στη λέξη w και $x|y$ σημαίνει ότι το x διαιρεί το y.

Άσκηση 30. Αν M είναι DFA δείξτε πώς να κατασκευάσετε ένα DFA για τη γλώσσα $\Sigma^* \setminus L(M)$ (το συμπλήρωμα της γλώσσας $L(M)$).

2.2 Μη ντετερμινιστικά αυτόματα

Εισάγουμε τώρα μια παραλλαγή των ντετερμινιστικών αυτομάτων, τα *μη ντετερμινιστικά αυτόματα* (Non-deterministic Finite Automata ή NFA). Αυτά αποτελούν, φαινομενικά, μια ενίσχυση του μοντέλου των ντετερμινιστικών αυτομάτων, αφού, από τον ορισμό που θα δώσουμε, κάθε DFA θα είναι και NFA, και άρα οι γλώσσες που αναγνωρίζονται από τα NFA είναι ένα υπερσύνολο των γλωσσών που αναγνωρίζονται από DFA.

Θα δούμε όμως σύντομα ότι αυτό είναι μόνο φαινομενικό, και ότι τα δύο μοντέλα είναι απολύτως ισοδύναμα, όσον αφορά τουλάχιστον το ποιες γλώσσες αναγνωρίζουν. Με άλλα λόγια, θα δούμε μια μέθοδο που για κάθε NFA θα κατασκευάζει ένα ισοδύναμο DFA, ένα DFA δηλ. που θα αναγνωρίζει ακριβώς την ίδια γλώσσα με το δοθέν NFA. Παρ΄ όλα αυτα το να μελετούμε NFA προσφέρει μερικά πλεονεκτήματα σε σχέση με τα DFA σε ορισμένες περιπτώσεις, όπως θα δούμε στα παραδείγματα.

Ορισμός 16. (Μη ντετερμινιστικό Αυτόματο (NFA))
Ένα *μη ντετερμινιστικό αυτόματο* είναι μια πεντάδα

$$(Q, \Sigma, \delta, q_0, F)$$

όπου

- Q είναι ένα πεπερασμένο σύνολο *καταστάσεων*,
- Σ είναι ένα πεπερασμένο αλφάβητο,
- δ είναι η *συνάρτηση μετάβασης* (transition function) με πεδίο ορισμού το $Q \times \Sigma$ και πεδίο τιμών το δυναμοσύνολο 2^Q, το σύνολο δηλ. όλων των υποσυνόλων του Q.
- $q_0 \in Q$ είναι μια από τις καταστάσεις που ονομάζεται *αρχική*, και
- $F \subseteq Q$ είναι το σύνολο των *τελικών* καταστάσεων.

Ένα μη ντετερμινιστικό αυτόματο είναι σαν ένα DFA αλλά για κάθε κορυφή και για κάθε γράμμα του αλφαβήτου μπορεί κανεις να έχει οποιοδήποτε πεπερασμένο πλήθος από ακμές που ξεκινούν από την κορυφή και έχουν το γράμμα ως ετικέτα Μπορεί ακόμη και να υπάρχει καμία ακμή από κάποια κορυφή για κάποιο γράμμα. (Αν $\delta(q, \alpha) = \emptyset$ είναι το κενό σύνολο τότε δεν υπάρχει στο γράφημα καμία ακμή από την κορυφή q με ετικέτα α.)

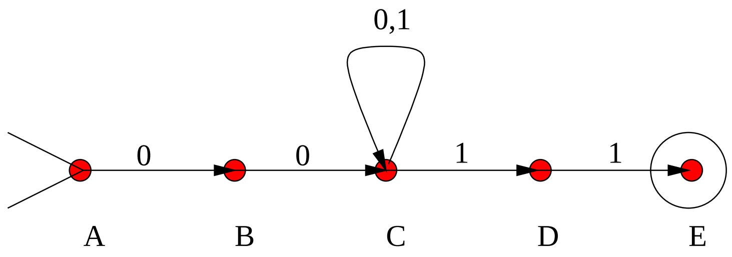
Η σημαντική διαφορά με πριν είναι ότι, όταν διαβάζουμε μια λέξη, δεν είναι πλέον μονοσήμαντα καθορισμένη η κίνησή μας πάνω στο γράφημα αφού ενδέχεται να έχουμε περισσότερες από μία επιλογές μετάβασης όντας σε μία κορυφή και διαβάζοντας ένα γράμμα. Το σε ποιες καταστάσεις μπορούμε να πάμε από μια κατάσταση $q \in Q$ αφού διαβάσουμε το γράμμα α μας το λέει η συνάρτηση μετάβασης. Το σύνολο αυτών των δυνατών καταστάσεων είναι το $\delta(q, \alpha) \subseteq Q$.

Παρατήρηση 4. Αν για μια κατάσταση q ενός NFA και για ένα γράμμα $\alpha \in \Sigma$ έχουμε $\delta(q, \alpha) = \emptyset$ αυτό ερμηνεύεται ως να μην είναι δυνατή η μετάβαση από την κατάσταση q προς οποιαδήποτε άλλη κατάσταση με το γράμμα α.

Ορισμός 17. (Αναγνώριση λέξης από NFA)
Θεωρούμε ότι μια λέξη $w = w_1 w_2 \ldots w_n$, $w_j \in \Sigma$, αναγνωρίζεται από ένα NFA αν υπάρχει τρόπος να κινηθούμε, διαβάζοντας το w, πάνω στο γράφημα ώστε να καταλήξουμε σε τελική κορυφή. Ξεκινούμε ευρισκόμενοι στην αρχική κατάσταση q_0 και διαβάζουμε τα γράμματα w_1 έως w_n με αυτή τη σειρά. Κάθε φορά που διαβάζουμε ένα γράμμα w_j, και ευρισκόμενοι στην κατάσταση q, επιλέγουμε ως επόμενη κατάσταση μια από τις καταστάσεις του συνόλου $\delta(q, w_j)$. (Αν το σύνολο αυτό είναι κενό η λέξη απορρίπτεται.) Είναι φανερό ότι για κάθε λέξη υπάρχουν ενδεχομένως περισσότεροι από ένας τρόποι να κινηθούμε πάνω στο αυτόματο διαβάζοντας τα γράμματα αυτής της λέξης, αφού σε κάθε βήμα ενδέχεται να έχουμε τη δυνατότητα να επιλέξουμε την επόμενή μας κατάσταση. Απορρίπτεται η λέξη w αν και μόνο αν κανείς από τους δυνατούς τρόπους κίνησης δεν καταλήγει σε τελική κορυφή.

Ορισμός 18. (Γλώσσα ενός NFA)
Το σύνολο των λέξεων που αποδέχεται το μη ντετερμινιστκό αυτόματο M ονομάζεται *η γλώσσα που αναγνωρίζει το αυτόματο* και συμβολίζεται με $L(M)$.

Παράδειγμα 27. Το ακόλουθο παράδειγμα μη ντετερμινιστικού αυτομάτου αναγνωρίζει τη γλώσσα $00\{0,1\}^*11$.



Σχήμα 7. NFA για τη γλώσσα $00\{0,1\}^*11$

Πρόκειται για NFA και όχι για DFA μια και στη μεσαία κορυφή υπάρχουν δύο ακμές με ετικέτα 1. Παρατηρείστε τη συντομογραφία που επιλέξαμε εδώ: ο βρόγχος (loop) από τη μεσαία κορυφή στον εαυτό της έχει δύο ετικέτες 0 και 1. Αυτό σημαίνει ότι επιτρέπεται να διανύσουμε την κορυφή αυτή όταν διαβάσουμε είτε 0 είτε 1. Θα μπορούσαμε να είχαμε το ίδιο αποτέλεσμα αν φτιάχναμε δύο βρόγχους από την κορυφή αυτή στον εαυτό της, ένα με ετικέτα 0 κι ένα με ετικέτα 1. Επιλέγουμε τη συντομογραφία για καθαρότητα στο σχήμα και τίποτε άλλο.

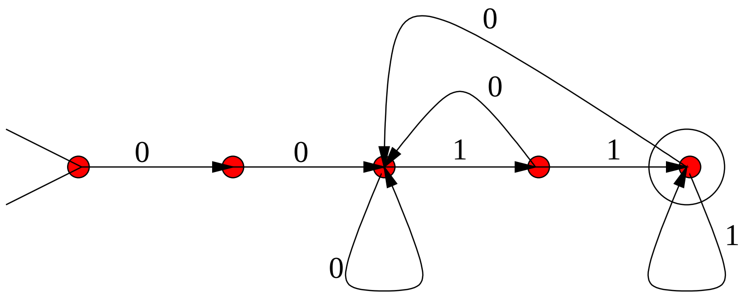
Γιατί το NFA του σχήματος 7 αναγνωρίζει τη γλώσσα $L = 00\{0,1\}^*11$;Παρατηρείστε ότι η L αποτελείται από εκείνες ακριβώς τις λέξεις που αρχίζουν με 00 και τελειώνουν με 11. Κάθε τέτοια λέξη περνάει από το NFA ως εξής: με τα δύο πρώτα 0 μεταβαίνει στη μεσαία κορυφή, στην οποία και παραμένει χρησιμοποιώντας το βρόγχο μέχρι να διαβάσει και το προπροτελευταίο γράμμα. Με τα δύο τελευταία 1 επιλέγει να κινηθεί προς τα δεξιά για να καταλήξει στη μοναδική τελική κορυφή. Θα μπορούσε να επιλέξει με τα δύο τελευταία 1 να κινηθεί πάνω στο βρόγχο της μεσαίας κορυφής, αλλά με αυτό τον τρόπο δεν καταλήγει σε τελική κορυφή. Αρκεί όμως για την αποδοχή μιας λέξης να υπάρχει έστω και ένας τρόπος να γίνει αποδεκτή.

Αντίστροφα τώρα, αν η λέξη w περνά από το NFA τότε υποχρεωτικά αρχίζει από δύο 0 (αλλιώς δε μπορεί να πάει δεξιότερα από τη δεύτερη κορυφή) και τελειώνει σε δύο 1 (αλλιώς δε μπορεί να καταλήξει στην τελική κορυφή).

Ας δώσουμε τώρα και ένα DFA για την ίδια γλώσσα:

Είναι μια καλή άσκηση να να αποδείξετε ότι το άνω αυτόματο αναγνωρίζει τη γλώσσα $00\{0,1\}^*11$. Για να το κάνετε δώστε κάποιο 'νόημα' στις καταστάσεις, όπως κάναμε παραπάνω για να δείξουμε ότι το DFA του σχήματος 5 δουλεύει.

Παρατήρηση 5. Συγκρίνοντας το NFA και το DFA για τη γλώσσα $00\{0,1\}^*11$ φαίνεται καθαρά το πλεονέκτημα του να σχεδιάζουμε NFA αντί για DFA. Επιτρέποντας στον εαυτό μας ένα πιο διευρυμένο μοντέλο μπορούμε να σχεδιάζουμε ευκολότερα αυτόματα για συγκεκριμένες γλώσσες, και η απλούστερη κατασκευή



Σχήμα 8. DFA για τη γλώσσα $00\{0,1\}^*11$

συνήθως επιτρέπει και μια ευκολότερη ανάλυση του γιατί το συγκεκριμένο αυτόματο δουλεύει. Και, θα δούμε αργότερα, δε χάνεται τίποτα σχεδιάζοντας NFA αντί για DFA, γιατί αυτά τα δύο υπολογιστικά μοντέλα είναι ισοδύναμα. Όχι μόνο αυτό, αλλά ο τρόπος μετατροπής ενός NFA σε DFA μπορεί να είναι τελείως μηχανικός (αλγοριθμικός).

Παράδειγμα 28. Στο αυτόματο που φαίνεται στο σχήμα 7 παραπάνω έχουμε $Q = \{A, B, C, D, E\}$, $\Sigma = \{0, 1\}$, $q_0 = A$, $F = \{E\}$, και

$$\begin{array}{rcl}
\delta(A,0) &=& \{B\}\\
\delta(A,1) &=& \emptyset\\
\delta(B,0) &=& \{C\}\\
\delta(B,1) &=& \emptyset\\
\delta(C,0) &=& \{C\}\\
\delta(C,1) &=& \{C,D\}\\
\delta(D,0) &=& \emptyset\\
\delta(D,1) &=& \{E\}\\
\delta(E,0) &=& \emptyset\\
\delta(E,1) &=& \emptyset
\end{array}$$

Άσκηση 31. Κατασκευάστε ένα NFA που να αναγνωρίζει τη γλώσσα της άσκησης 25.

Άσκηση 32. Κατασκευάστε ένα NFA που να αναγνωρίζει εκείνες τις λέξεις πάνω από το $\Sigma = \{0, 1\}$ που περιέχουν δύο μηδενικά που απέχουν μεταξύ τους στη λέξη κατά πολλαπλάσιο του 4. Δύο διαδοχικά μηδενικά θεωρούνται ότι απέχουν απόσταση μηδέν, άρα πολλαπλάσιο του 4.

2.3 Ισοδυναμία NFA και DFA.

Επαναλαμβάνουμε ότι για ένα NFA η συνάρτηση μετάβασης δ είναι ο τρόπος κωδικοποίησης των μεταβάσεων του. Αν δηλ. $q \in Q$ είναι μια κατάσταση και $\alpha \in \Sigma$ είναι ένα γράμμα του αλφαβήτου, το σύνολο $\delta(q, \alpha) \subseteq Q$ είναι το σύνολο όλων των καταστάσεων του NFA στις οποίες μπορεί αυτό να μεταβεί όντας στην κατάσταση q και διαβάζοντας το γράμμα α.

Ορισμός 19. (Η συνάρτηση δ πάνω σε σύνολα)
Αν $A \subseteq Q$ είναι ένα σύνολο καταστάσεων και $B \subseteq \Sigma$ είναι ένα σύνολο γραμμάτων τότε

$$\delta(A, B) = \bigcup_{q \in A, \alpha \in B} \delta(q, \alpha).$$

Δηλ. $\delta(A,B)$ είναι το σύνολο όλων των καταστάσεων στις οποίες μπορεί κανείς να φτάσει ξεκινώντας από κάποια κατάσταση του A και ακολουθώντας κάποιο γράμμα α του συνόλου B. Η επέκταση αυτή της συνάρτησης δ ώστε να παίρνει ως όρισμα σύνολα αντί για γράμματα είναι η συνηθισμένη επεκταση μιας συνάρτησης $f : C \to D$ που με $f(X)$, $X \subseteq C$, συμβολίζει το σύνολο εικόνων του συνόλου X μέσω της f.

Επεκτείνουμε τώρα το πεδίο ορισμού της συνάρτησης δ όσον αφορά το δεύτερο όρισμά της.

Ορισμός 20. (Η συνάρτηση μετάβασης δ πάνω σε λέξεις)
Αν $q \in Q$ είναι μια κατάσταση και $w \in \Sigma^*$ είναι μια λέξη (ενδεχομένως και κενή) τότε ορίζουμε το σύνολο καταστάσεων $\delta(q,w)$ ως εξής:

$$\delta(q,w) = \begin{cases} \{q\} & \epsilon\alpha\nu \quad w = \epsilon \\ \delta(\delta(q,v),\alpha) & \epsilon\alpha\nu \quad w = v\alpha, v \in \Sigma^*, \alpha \in \Sigma. \end{cases} \tag{2.1}$$

Παρατήρηση 6. Ο ορισμός (2.1) απαιτεί λίγη προσοχή όσον αφορά το δεύτερο σκέλος του μια και είναι *αυτοαναφορικός*, αναφερόμαστε δηλ. στη συνάρτηση δ για να ορίσουμε τη συνάρτηση δ. Η ουσία εδώ είναι ότι αυτή η αυτοαναφορά δε δημιουργεί κάποιο πρόβλημα, μια και για να ορίσουμε το $\delta(q,w)$ αναφερόμαστε σε τιμές του $\delta(q,x)$ για λέξεις x με μήκος $|x| < |w|$, και υπάρχει και ο ξεχωριστός ορισμός για τη λέξη μήκους 0 όπου σταματά η αυτοαναφορά. Για να υπολογιστεί έτσι το $\delta(q,w)$ υπολογίζεται πρώτα το δ για ολοένα και μικρότερες λέξεις, ώσπου τελικά φτάνει να χρησιμοποιείται το πρώτο μέλος του ορισμού (2.1) που δεν έχει αυτοαναφορά. Για παράδειγμα, αν θέλει κανείς να υπολογίσει την τιμή $\delta(q,abc)$, όπου $a,b,c \in \Sigma$, ακολουθώντας τον ορισμό (2.1) αυτό γίνεται ως εξής:

$$\begin{aligned} \delta(q,abc) &= \delta(\delta(q,ab),c) \\ &= \delta(\delta(\delta(q,a),b),c). \end{aligned}$$

Παρατηρούμε ότι στην τελευταία γραμμή η συνάρτηση δ είναι η γνωστή μας συνάρτηση που σαν δεύτερο όρισμα έχει σύμβολο του αλφαβήτου και όχι λέξη. Ορισμοί όπως ο (2.1) είναι πολύ κοινοί και ονομάζονται *αναδρομικοί*.

Με άλλα λόγια $\delta(q,w), w \in \Sigma^*$, είναι το σύνολο όλων εκείνων των καταστάσεων του αυτομάτου στις οποίες μπορεί κανείς να βρεθεί ξεκινώντας από την κατάσταση q και ακολουθώντας τη λέξη w. Και, αν $R \subseteq Q$ είναι ένα σύνολο καταστάσεων, με $\delta(R,w)$ συμβολίζεται το σύνολο όλων των καταστάσεων στις οποίες μπορεί κανείς να πάει ξεκινώντας από κάποια κατάσταση στο R και ακολουθώντας τη λέξη w.

Παρατήρηση 7. Μια λέξη w αναγνωρίζεται από ένα αυτόματο αν και μόνο αν

$$\delta(q_0,w) \cap F \neq \emptyset,$$

όπου F το σύνολο των τελικών καταστάσεων του NFA. Αναγνωρίζεται δηλ. μια λέξη αν και μόνο αν *είναι δυνατή η μετάβαση* από την αρχική σε *κάποια* τελική κατάσταση ακολουθώντας τη λέξη.

Ορισμός 21. (Ισοδυναμία αυτομάτων)
Δύο ντετερμινιστικά ή μη ντετερμινιστικά αυτόματα M_1 και M_2 λέγονται *ισοδύναμα* αν και μόνο αν $L(M_1) = L(M_2)$, αν αναγνωρίζουν δηλ. ακριβώς τις ίδιες λέξεις.

Θεώρημα 3. (Ισοδυναμία NFA και DFA)
Για κάθε NFA M υπάρχει ισοδύναμο DFA M'.

Απόδειξη. Περιγράφουμε ένα αλγόριθμο μετάβασης από ένα τυχόν NFA M σε ένα *ισοδύναμο* DFA M'.

Έστω λοιπόν ότι το NFA M είναι η πεντάδα (Q,Σ,δ,q_0,F). Το DFA $M' = (Q',\Sigma,\delta',q_0',F')$ θα έχει σύνολο καταστάσεων $Q' = 2^Q$ το δυναμοσύνολο (σύνολο όλων των υποσυνόλων) του Q, αρχική κατάσταση $q_0' = \{q_0\}$, ίδιο αλφάβητο Σ και τελικές καταστάσεις όλα εκείνα τα σύνολα καταστάσεων του M που περιέχουν κάποια τελική κατάσταση

$$F' = \{A \subseteq Q : A \cap F \neq \emptyset\}.$$

Τέλος η συνάρτηση μετάβασης $\delta' : Q' \times \Sigma \to Q'$ του M' ορίζεται ως

$$\delta'(q',\alpha) = \delta(q',\alpha).$$

Θυμηθείτε ότι το σύμβολο q' παριστάνει μια κατάσταση του M' άρα ένα σύνολο καταστάσεων του M, και άρα η δ συνάρτηση που εμφανίζεται δεξιά στον άνω ορισμό είναι η επεκτεταμένη δ συνάρτηση του αυτομάτου M όπως την ορίσαμε παραπάνω, μια και σαν πρώτο όρισμα έχει ολόκληρο σύνολο.

Με άλλα λόγια: στο DFA M' που κατασκευάζουμε από την κατάσταση q_1 (υποσύνολο του Q) με το σύμβολο $\alpha \in \Sigma$ μεταβαίνουμε στην κατάσταση q_2, που είναι το σύνολο όλων εκείνων των καταστάσεων του M στις οποίες μπορούμε να μεταβούμε από κάποια κατάσταση του συνόλου q_1 με το γράμμα α κινούμενοι πάνω στο M. Δείτε το Παράδειγμα 29 παρακάτω.

Άσκηση 33. Αποδείξτε με επαγωγή ως προς το μήκος $|w|$ της λέξης $w \in \Sigma^*$ ότι για κάθε κατάσταση q του M' ισχύει

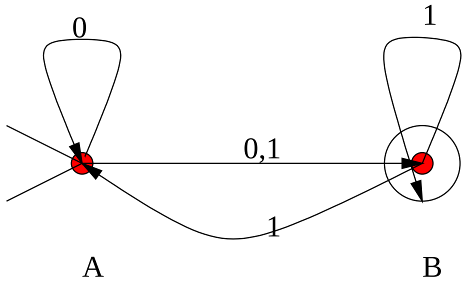
$$\delta'(q, w) = \delta(q, w).$$

Για να δείξουμε το Θεώρημα αρκεί να δείξουμε την ισοδυναμία

$$w \in L(M) \Longleftrightarrow w \in L(M') \tag{2.2}$$

για όλες τις λέξεις $w \in \Sigma^*$. Αλλά $w \in L(M)$ αν και μόνο αν $\delta(q_0, w) \in F$, ενώ $w \in L(M')$ αν και μόνο αν $\delta'(q_0', w) \in F'$. Χρησιμοποιώντας το Πρόβλημα 33 το τελευταίο συμβαίνει αν και μόνο αν $\delta(\{q_0\}, w) \in F'$ και, χρησιμοποιώντας τον ορισμό του F' βλέπυμε ότι αυτό ισχύει αν και μόνο αν $\delta(q_0, w) \cap F \neq \emptyset$. Όμως το σύνολο $\delta(q_0, w)$ είναι μονοσύνολο μια και το M είναι DFA, άρα η τελευταία συνθήκη είναι ισοδύναμη με $\delta(q_0, w) \in F$, και η απόδειξη είναι πλήρης.
□

Παράδειγμα 29. Ας δούμε τώρα πώς μετατρέπεται το ακόλουθο απλό NFA του Σχήματος 9 σε DFA. Το αποτέλεσμα δίνεται στο Σχήμα 10.



Σχήμα 9. M, ένα απλό NFA

Το σύνολο καταστάσεων του M' θα είναι το δυναμοσύνολο του συνόλου καταστάσεων του M δηλ. του $\{A, B\}$. Είναι δηλ. το σύνολο

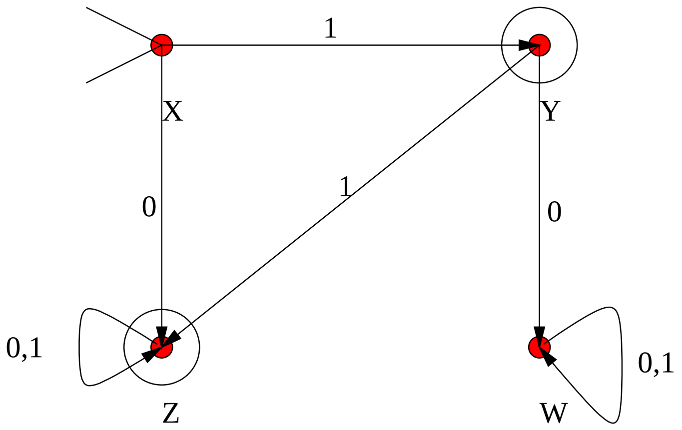
$$Q' = \{X = \{A\}, Y = \{B\}, Z = \{A, B\}, W = \emptyset\}.$$

Σύνολο τελικών καταστασεων είναι το $F' = \{Y, Z\}$, όλα εκείνα τα σύνολα δηλ. που περιέχουν κάποια τελική κατάσταση του M. Αρχική κατάσταση είναι η X.

Άσκηση 34. Ελέγξτε την ορθότητα την ισοδυναμίας του NFA του σχήματος 9 στο DFA του σχήματος 10, πάνω στις λέξεις 001 και 10.

Άσκηση 35. Φτιάξτε ισοδύναμα DFA για τα NFA

1. $(Q = \{p, q, r, s\}, \Sigma = \{0, 1\}, \delta_1, q_0 = p, F = \{s\})$
2. $(Q = \{p, q, r, s\}, \Sigma = \{0, 1\}, \delta_2, q_0 = p, F = \{q, s\})$



Σχήμα 10. Το NFA M του Σχήματος 9 μετατετραμμένο στο DFA M'

όπου οι συναρτήσεις μετάβασης είναι οι

$$\delta_1 : \begin{array}{c|r|r} & 0 & 1 \\ \hline p & \{p,q\} & \{p\} \\ q & \{r\} & \{r\} \\ r & \{s\} & \emptyset \\ s & \{s\} & \{s\} \end{array} \qquad \delta_2 : \begin{array}{c|r|r} & 0 & 1 \\ \hline p & \{q,s\} & \{q\} \\ q & \{r\} & \{q,r\} \\ r & \{s\} & \{p\} \\ s & \emptyset & \{p\} \end{array}$$

Άσκηση 36. Αν ένα NFA έχει n το πλήθος καταστάσεις, πόσες το πολύ καταστάσεις χρειάζεται να έχει ένα ισοδύναμό του DFA;

2.4 NFA με ϵ-κινήσεις

Μια ακόμη παραλλαγή του μοντέλου του πεπερασμένου αυτομάτου είναι το μη ντετερμινιστικό αυτόματο με ϵ-κινήσεις. Αυτό είναι ένα NFA που έχει, ενδεχομένως, και κάποιες ακμές που αντί να έχουν ως ετικέτα ένα γράμμα του αλφαβήτου έχουν την κενή λέξη ϵ. Κινούμενοι πάνω στο αυτόματο αυτό για να αναγνωρίσουμε μια λέξη έχουμε το δικαίωμα να διανύσουμε μια ϵ-ακμή οποτεδήποτε υπάρχει μια τέτοια από την τρέχουσα κατάσταση *χωρίς να μας ενδιαφέρει ποιο είναι το επόμενο γράμμα της λέξης*. Η διάνυση μιας ϵ-ακμής *δε συνοδεύεται από μεταπήδηση στο επόμενο γράμμα της λέξης*. Ας τα λέμε αυτά ϵ-NFA.

Ορισμός 22. (Μη ντετερμινιστικό Αυτόματο με ϵ-κινήσεις (ϵ-NFA))
Ένα *μη ντετερμινιστικό αυτόματο* με ϵ-κινήσεις (ϵ-NFA) είναι μια πεντάδα

$$(Q, \Sigma, \delta, q_0, F)$$

όπου

- Q είναι ένα πεπερασμένο σύνολο *καταστάσεων*,
- Σ είναι ένα πεπερασμένο αλφάβητο,
- δ είναι η *συνάρτηση μετάβασης* (transition function) με πεδίο ορισμού το $Q \times (\Sigma \cup \{\epsilon\})$ και πεδίο τιμών το δυναμοσύνολο 2^Q, το σύνολο δηλ. όλων των υποσυνόλων του Q.
- $q_0 \in Q$ είναι μια από τις καταστάσεις που ονομάζεται *αρχική*, και
- $F \subseteq Q$ είναι το σύνολο των *τελικών* καταστάσεων.

Παρατήρηση 8. Η ουσιαστική διαφορά από τον ορισμό του NFA είναι ότι η συνάρτηση μετάβασης μπορεί να έχει και το ϵ ως δεύτερο όρισμα και όχι απαραίτητα ένα γράμμα του αλφαβήτου Σ.

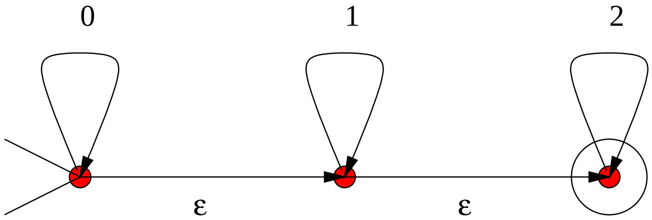
Παρατήρηση 9. Το νόημα μιας ϵ-ακμής από μια κορυφή q_1 σε μια κορυφή q_2 είναι το εξής: αν το NFA βρίσκεται στην κατάσταση q_1 τότε μπορεί να επιλέξει να μεταβεί στην κατάσταση q_2 *χωρίς να διαβάσει το επόμενο γράμμα της λέξης*.

Ορισμός 23. (Αναγνώριση λέξης από ϵ-NFA)
Θεωρούμε ότι μια λέξη $w = w_1 w_2 \ldots w_n$, $w_j \in \Sigma$, αναγνωρίζεται από ένα ϵ-NFA αν υπάρχει τρόπος να κινηθούμε, διαβάζοντας το w, πάνω στο γράφημα ώστε να καταλήξουμε σε τελική κορυφή. Ξεκινούμε ευρισκόμενοι στην αρχική κατάσταση q_0 και διαβάζουμε τα γράμματα w_1 έως w_n με αυτή τη σειρά. Σε κάθε βήμα έχουμε τη δυνατότητα να διαβάσουμε το επόμενο γράμμα w_j και να εκτελέσουμε μια μετάβαση σε νέα κατάσταση ή όχι.

Κάθε φορά που διαβάζουμε ένα γράμμα w_j, και ευρισκόμενοι στην κατάσταση q, επιλέγουμε ως επόμενη κατάσταση μια από τις καταστάσεις του συνόλου $\delta(q, w_j)$. (Αν το σύνολο αυτό είναι κενό η λέξη απορρίπτεται.) Αν επιλέξουμε να μη διαβάσουμε το επόμενο γράμμα της λέξης μπορούμε να μεταβούμε σε νέα κατάσταση ακολουθώντας μια ϵ-ακμή, μπορούμε δηλ. να μεταβούμε σε μια αποιαδήποτε από τις καταστάσεις του συνόλου $\delta(q, \epsilon)$, όπου q είναι η τρέχουσα κατάσταση. (Αν το σύνολο αυτό είναι κενό η λέξη απορρίπτεται.) Είναι φανερό ότι για κάθε λέξη υπάρχουν ενδεχομένως περισσότεροι από ένας τρόποι να κινηθούμε πάνω στο αυτόματο διαβάζοντας τα γράμματα αυτής της λέξης, αφού σε κάθε βήμα ενδέχεται να έχουμε τη δυνατότητα να επιλέξουμε την επόμενή μας κατάσταση. Απορρίπτεται η λέξη w αν και μόνο αν κανείς από τους δυνατούς τρόπους κίνησης δεν καταλήγει σε τελική κορυφή.

Ορισμός 24. (Γλώσσα ενός ϵ-NFA)
Το σύνολο των λέξεων που αποδέχεται το μη ντετερμινιστικό αυτόματο με ϵ κινήσεις M ονομάζεται *η γλώσσα που αναγνωρίζει το αυτόματο* και συμβολίζεται με $L(M)$.

Παράδειγμα 30. Το ϵ-NFA του Σχήματος 11 αναγνωρίζει τη γλώσσα $0^*1^*2^*$. Για παράδειγμα, πώς αναγνω-



Σχήμα 11. Ένα ϵ-NFA για τη γλώσσα $0^*1^*2^*$.

ρίζει τη λέξη $0^2 1^3 2^5$ (συντομογραφία για τη λέξη 0011122222); Κάνει δύο μεταβάσεις από την πρώτη κορυφή στον εαυτό της διαβάζοντας τα δύο 0, μετά μεταβαίνει στη δεύτερη κορυφή *χωρίς να διαβάσει τίποτα*, μεταβαίνει από τη δεύτερη κορυφή στον εαυτό της τρεις φορές διαβάζοντας τα 1, μεταβαίνει στην τρίτη κορυφή χωρίς πάλι να διαβάσει τίποτα, και τέλος μεταβαίνει από την τρίτη κορυφή στον εαυτό της διαβάζοντας τα πέντε 2.

Άσκηση 37. Κατασκευάστε ένα ϵ-NFA που να αναγνωρίζει τη γλώσσα $(0101)^* \cup (111)^*$.

Άσκηση 38. Αν M_1 και M_2 είναι δύο DFA περιγράψτε πως θα τα χρησιμοποιήσετε αυτά για να κατασκευάσετε ένα ϵ-NFA για τη γλώσσα $L(M_1) \cup L(M_2)$.

Άσκηση 39. Αν M_1 και M_2 είναι δύο DFA περιγράψτε πως θα τα χρησιμοποιήσετε αυτά για να κατασκευάσετε ένα ϵ-NFA για τη γλώσσα $L(M_1)L(M_2)$. Αν το M_1 έχει ακριβώς μια τελική κατάσταση κατασκευάστε DFA για τη γλώσσα αυτή.

2.5 Ισοδυναμία ϵ-NFA και NFA

Εδώ θα αποδείξουμε ότι τα ϵ-NFA είναι ισοδύναμα με τα NFA, και άρα και με τα DFA. Για κάθε ϵ-NFA δηλ. υπάρχει ένα NFA χωρίς ϵ-κινήσεις που αναγνωρίζει την ίδια γλώσσα.

Ορισμός 25. (ϵ-μονοπάτι)
Ένα *ϵ-μονοπάτι* στο γράφημα ενός ϵ-NFA είναι μια πεπερασμένη ακολουθία διαδοχικών ϵ-ακμών (ακμών δηλ. που φέρουν την ετικέτα ϵ).

Ορισμός 26. (α-μονοπάτι)
Ένα *α-μονοπάτι*, όπου $\alpha \in \Sigma$, είναι ένα μονοπάτι στο γράφημα ενός ϵ-NFA που απαρτίζεται από ένα αρχικό ϵ-μονοπάτι, ακολουθούμενο από μια ακμή με ετικέτα α, ακολουθούμενο από ένα ϵ-μονοπάτι. Τα δύο ϵ-μονοπάτια μπορούν να είναι και κενά.

Θεώρημα 4. (Ισοδυναμία ϵ-NFA και NFA)
Για κάθε ϵ-NFA M υπάρχει ισοδύναμο NFA M'.

Απόδειξη. Για να μεταβούμε από ένα ϵ-NFA M σε ένα ισοδύναμο NFA M' κάνουμε τα εξής:

- Το σύνολο καταστάσεων του M' είναι ίδιο με του M.
- Η αρχική κατάσταση του M' είναι ίδια με αυτή του M.
- Οι τελικές καταστάσεις του M', το σύνολο των οποίων συμβολίζουμε με F', αποτελείται από όλες εκείνες τις καταστάσεις στις οποίες μπορούμε να φτάσουμε από κάποια κορυφή του F (το σύνολο τελικών καταστάσεων του M) με ένα ϵ-μονοπάτι. Επειδή ένα κενό μονοπάτι είναι ϵ-μονοπάτι, ο ορισμός αυτός συνεπάγεται ότι $F \subseteq F'$.
- Για κάθε κατάσταση q και γράμμα $\alpha \in \Sigma$ θέτουμε στο M' ως $\delta_{M'}(q, \alpha)$ το σύνολο όλων των καταστάσεων στις οποίες μπορεί κανείς να μεταβεί από την q στο M διανύοντας κάποιο α-μονοπάτι.
- Καταργούμε όλες τις ϵ-κινήσεις.

Έστω τώρα $w = w_1 w_2 \cdots w_n \in \Sigma^*$, $n \geq 0$. Πρέπει να δείξουμε ότι η λέξη w γίνεται δεκτή από το M αν και μόνο αν γίνεται δεκτή από το M.

Αν $w \in L(M)$ αυτό σημαίνει ότι υπάρχει τρόπος κίνησης πάνω στο M, διαβάζοντας τα γράμματα της λέξης w είτε ακολουθώντας ϵ-κινήσεις, ώστε να μεταβούμε από την q_0 σε κάποια κατάσταση $f \in F$. Σε αυτή την περίπτωση δείχνουμε ότι υπάρχει τρόπος κίνησης πάνω στο M τέτοιος που να οδηγεί από την q_0 σε κάποια κατάσταση του F'.

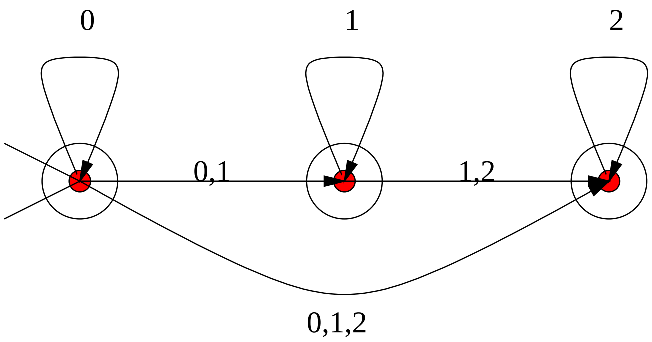
Έστω λοιπόν ότι $w \in L(M)$. Αυτό σημαίνει ότι υπάρχει μια ακολουθία καταστάσεων του M έστω $q_0, q_1, \ldots, q_{k-1}, q_k$, $k \geq 0$, τέτοια ώστε $q_k \in F$ και η μετάβαση από την q_j στην q_{j+1}, $0 \leq j < k$, γίνεται είτε με κάποιο γράμμα w_j είτε με μια ϵ-ακμή, και τα γράμματα w_j χρησιμοποιούνται όλα, από μια φορά το καθένα, και με τη σειρά.

Χωρίζουμε την πεπερασμένη αυτή ακολουθία των μεταβάσεων $q_0 \to q_1, q_1 \to q_2, \ldots, q_{k-1} \to q_k$ σε κομμάτια $q_{a_0} \to q_{a_1}, q_{a_1} \to q_{a_2}, \ldots, q_{a_{n-1}} \to q_{a_n}$ (με $a_0 = 0$, $a_n = k$) τα οποία αντιστοιχούν σε ένα w_1-μονοπάτι, ακολουθούμενο από ένα w_2-μονοπάτι, κλπ., τελειώνοντας με ένα w_n-μονοπάτι. Από τον ορισμό του NFA M' προκύπτει ότι στο M' είναι δυνατή η μετάβαση από την κορυφή q_{a_0} στην q_{a_1} με μια w_1-ακμή, από την q_{a_1} στην q_{a_2} με μια w_2-ακμή, κλπ. Επειδή η κατάσταση q_k παραμένει τελική κατάσταση και του M' προκύπτει ότι $w \in L(M')$.

Ο εγκλεισμός $L(M') \subseteq L(M)$ αφήνεται ως άσκηση (Άσκηση 40).
□

Άσκηση 40. Με το συμβολισμό της άνω απόδειξης δείξτε ότι αν μια λέξη αναγνωρίζεται από το NFA M' τότε αναγνωρίζεται και από το ϵ-NFA M.

Παράδειγμα 31. Με τη μέθοδο της απόδειξης του Θεωρήματος 4 το ϵ-NFA του Σχήματος 11 μετατρέπεται στο NFA του Σχήματος 12.



Σχήμα 12. Το ϵ-NFA του Σχήματος 11 μετατετραμμένο σε NFA