Πεπερασμένη περιγραφή (αναπαράσταση) γλωσσών

Μπορούμε να σκεφτούμε πολλές περιγραφές γλωσσών. Π.χ. μπορούμε να ορίσουμε ως εξής μια γλώσσα:

$$L = \{w : w \in \{0,1\}^* \text{ και περιέχει ίσο αριθμό από } 0 \text{ και } 1\}$$

Πεπερασμένη περιγραφή (αναπαράσταση) γλωσσών

Μπορούμε να σκεφτούμε πολλές περιγραφές γλωσσών. Π.χ. μπορούμε να ορίσουμε ως εξής μια γλώσσα:

$$L = \{w : w \in \{0,1\}^* \text{ και περιέχει ίσο αριθμό από } 0 \text{ και } 1\}$$

Γενικότερα οι περιγραφές είναι της μορφής:

$$L = \{w : w \text{ ικανοποιεί την ιδιότητα } T\}, \text{ για κάποιο } T$$

Πεπερασμένη περιγραφή (αναπαράσταση) γλωσσών

Μπορούμε να σκεφτούμε πολλές περιγραφές γλωσσών. Π.χ. μπορούμε να ορίσουμε ως εξής μια γλώσσα:

$$L = \{w : w \in \{0,1\}^* \text{ και περιέχει ίσο αριθμό από } 0 \text{ και } 1\}$$

Γενικότερα οι περιγραφές είναι της μορφής:

$$L = \{w : w \text{ ικανοποιεί την ιδιότητα } T\}, \text{ για κάποιο } T$$

Η ιδιότητα T πρέπει να έχει πεπερασμένη περιγραφή Π_T. Έστω Π το σύνολο αυτών των περιγραφών

$$\Pi = \{\Pi_T : T \text{ είναι ιδιότητα}\}$$

Πεπερασμένη περιγραφή (αναπαράσταση) γλωσσών

Μπορούμε να σκεφτούμε πολλές περιγραφές γλωσσών. Π.χ. μπορούμε να ορίσουμε ως εξής μια γλώσσα:

$$L = \{w : w \in \{0,1\}^* \text{ και περιέχει ίσο αριθμό από } 0 \text{ και } 1\}$$

Γενικότερα οι περιγραφές είναι της μορφής:

$$L = \{w : w \text{ ικανοποιεί την ιδιότητα } T\}, \text{ για κάποιο } T$$

Η ιδιότητα T πρέπει να έχει πεπερασμένη περιγραφή Π_T. Έστω Π το σύνολο αυτών των περιγραφών

$$\Pi = \{\Pi_T : T \text{ είναι ιδιότητα}\}$$

Άρα το Π είναι μία γλώσσα!

Πεπερασμένη περιγραφή (αναπαράσταση) γλωσσών

Μπορούμε να σκεφτούμε πολλές περιγραφές γλωσσών. Π.χ. μπορούμε να ορίσουμε ως εξής μια γλώσσα:

$$L = \{w : w \in \{0,1\}^* \text{ και περιέχει ίσο αριθμό από } 0 \text{ και } 1\}$$

Γενικότερα οι περιγραφές είναι της μορφής:

$$L = \{w : w \text{ ικανοποιεί την ιδιότητα } T\}, \text{ για κάποιο } T$$

Η ιδιότητα T πρέπει να έχει πεπερασμένη περιγραφή Π_T. Έστω Π το σύνολο αυτών των περιγραφών

$$\Pi = \{\Pi_T : T \text{ είναι ιδιότητα}\}$$

Άρα το Π είναι μία γλώσσα!

Γενικά υπάρχουν πολλά σύνολα–είδη περιγραφών. Εδώ θα ξεκινήσουμε με μια απλή και φυσική περιγραφή και στη συνέχεια θα την εμπλουτίσουμε.

Κανονικές εκφράσεις (ΚΕ / **Regular expressions**)

1. $\emptyset, \sigma \in \Sigma$

Κανονικές εκφράσεις (ΚΕ / **Regular expressions**)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ

Κανονικές εκφράσεις (ΚΕ / **Regular expressions**)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Γλώσσες και κανονικές εκφράσεις

Σε κάθε ΚΕ αντιστοιχούμε μία γλώσσα ως εξής: Αν α είναι μία ΚΕ, τότε $\mathcal{L}(\alpha)$ θα συμβολίζει τη γλώσσα.

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Γλώσσες και κανονικές εκφράσεις

Σε κάθε ΚΕ αντιστοιχούμε μία γλώσσα ως εξής: Αν α είναι μία ΚΕ, τότε $\mathcal{L}(\alpha)$ θα συμβολίζει τη γλώσσα.

1. $\mathcal{L}(\emptyset) = \emptyset, \mathcal{L}(\sigma) = \{\sigma\}$

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Γλώσσες και κανονικές εκφράσεις

Σε κάθε ΚΕ αντιστοιχούμε μία γλώσσα ως εξής: Αν α είναι μία ΚΕ, τότε $\mathcal{L}(\alpha)$ θα συμβολίζει τη γλώσσα.

1. $\mathcal{L}(\emptyset) = \emptyset, \mathcal{L}(\sigma) = \{\sigma\}$
2. $\mathcal{L}((\alpha \cup \beta)) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Γλώσσες και κανονικές εκφράσεις

Σε κάθε ΚΕ αντιστοιχούμε μία γλώσσα ως εξής: Αν α είναι μία ΚΕ, τότε $\mathcal{L}(\alpha)$ θα συμβολίζει τη γλώσσα.

1. $\mathcal{L}(\emptyset) = \emptyset, \mathcal{L}(\sigma) = \{\sigma\}$
2. $\mathcal{L}((\alpha \cup \beta)) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$

3. $\mathcal{L}((\alpha\beta)) = \mathcal{L}(\alpha)\mathcal{L}(\beta)$

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Γλώσσες και κανονικές εκφράσεις

Σε κάθε ΚΕ αντιστοιχούμε μία γλώσσα ως εξής: Αν α είναι μία ΚΕ, τότε $\mathcal{L}(\alpha)$ θα συμβολίζει τη γλώσσα.

1. $\mathcal{L}(\emptyset) = \emptyset, \mathcal{L}(\sigma) = \{\sigma\}$
2. $\mathcal{L}((\alpha \cup \beta)) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$
3. $\mathcal{L}((\alpha\beta)) = \mathcal{L}(\alpha)\mathcal{L}(\beta)$

4. $\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$

Μια γλώσσα L λεγεται κανονική αν και μόνο αν υπάρχει αντίστοιχη κανονική έκφραση ρ, δηλαδη $L = \mathcal{L}(\rho)$.

Παραδείγματα γλωσσών κανονικών εκφράσεων

Ποια είναι η γλώσσα της κανονικής έκφρασης $((a \cup b)^*)$;

Απαντηση:

Παραδείγματα γλωσσών κανονικών εκφράσεων

Ποια είναι η γλώσσα της κανονικής έκφρασης $((a \cup b)^*)$;
Απαντηση:

$$\begin{aligned}\mathcal{L}\left(((a \cup b)^*)\right) &= \mathcal{L}\left((a \cup b)\right)^* && [\text{κανονας } 4] \\ &= (\mathcal{L}(a) \cup \mathcal{L}(b))^* && [\text{κανονας } 2] \\ &= (\{a\} \cup \{b\})^* && [\text{κανονας } 1] \\ &= \{a, b\}^*\end{aligned}$$

Δηλαδή η γλώσσα που περιέχει όλες τις συμβολοσειρές.

Παραδείγματα γλωσσών κανονικών εκφράσεων

Ποια είναι η γλώσσα της κανονικής έκφρασης $((a \cup ba)^*)$;
Απαντηση:

Παραδείγματα γλωσσών κανονικών εκφράσεων

Ποια είναι η γλώσσα της κανονικής έκφρασης $((a \cup ba)^*)$;
Απαντηση:

$$\begin{aligned}\mathcal{L}(((a \cup (ba))^*)) &= \mathcal{L}((a \cup (ba)))^* && [\text{κανονας } 4] \\ &= (\mathcal{L}(a) \cup \mathcal{L}((ba)))^* && [\text{κανονας } 2] \\ &= (\mathcal{L}(a) \cup \mathcal{L}(b)\mathcal{L}(a)))^* && [\text{κανονας } 3] \\ &= (\{a\} \cup \{ba\})^* && [\text{κανονας } 1] \\ &= \{a, ba\}^*\end{aligned}$$

Παραδείγματα γλωσσών κανονικών εκφράσεων

Ποια είναι η γλώσσα της κανονικής έκφρασης $((a \cup ba)^*)$;
Απαντηση:

$$\begin{aligned}\mathcal{L}(((a \cup (ba))^*)) &= \mathcal{L}((a \cup (ba)))^* && [\text{κανονας } 4] \\ &= (\mathcal{L}(a) \cup \mathcal{L}((ba)))^* && [\text{κανονας } 2] \\ &= (\mathcal{L}(a) \cup \mathcal{L}(b)\mathcal{L}(a)))^* && [\text{κανονας } 3] \\ &= (\{a\} \cup \{ba\})^* && [\text{κανονας } 1] \\ &= \{a, ba\}^*\end{aligned}$$

Δηλαδή η γλώσσα που περιέχει τις συμβολοσειρές στις οποίες κάθε b ακολουθείται πάντα απο a.

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Κανονικές εκφράσεις (ΚΕ / Regular expressions)

1. $\emptyset, \sigma \in \Sigma$
2. Αν α και β είναι ΚΕ, τότε $(\alpha \cup \beta)$ είναι ΚΕ
3. Αν α και β είναι ΚΕ, τότε $(\alpha\beta)$ είναι ΚΕ
4. Αν α είναι ΚΕ, τότε (α^*) είναι ΚΕ
5. Τίποτα άλλο δεν είναι ΚΕ

Γλώσσες και κανονικές εκφράσεις

Σε κάθε ΚΕ αντιστοιχούμε μία γλώσσα ως εξής: Αν α είναι μία ΚΕ, τότε $\mathcal{L}(\alpha)$ θα συμβολίζει τη γλώσσα.

1. $\mathcal{L}(\emptyset) = \emptyset, \mathcal{L}(\sigma) = \{\sigma\}$
2. $\mathcal{L}((\alpha \cup \beta)) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$
3. $\mathcal{L}((\alpha\beta)) = \mathcal{L}(\alpha)\mathcal{L}(\beta)$
4. $\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$

Μια γλώσσα L λεγεται κανονική αν και μόνο αν υπάρχει αντίστοιχη κανονική έκφραση ρ, δηλαδη $L = \mathcal{L}(\rho)$.

Παραδείγματα γλωσσών κανονικών εκφράσεων

Δώστε μια κανονική έκφραση για τη γλώσσα

$$L = \{w : w \in \{a, b\}^* \text{ και } w \text{ περιεχει περιττό αριθμό από } a\}.$$

Απαντηση:

Παραδείγματα γλωσσών κανονικών εκφράσεων

Δώστε μια κανονική έκφραση για τη γλώσσα

$$L = \{w : w \in \{a, b\}^* \text{ και } w \text{ περιεχει περιττό αριθμό από } a\}.$$

Απαντηση:

$b^*a(b^*ab^*a)^*b^*$ όπου παραλείπονται αυτονόητες παρενθέσεις.

Παραδείγματα γλωσσών κανονικών εκφράσεων

Δώστε μια κανονική έκφραση για τη γλώσσα

$$L = \{w : w \in \{a, b\}^* \text{ και } w \text{ περιεχει περιττό αριθμό από } a\}.$$

Απαντηση:
$b^*a(b^*ab^*a)^*b^*$ όπου παραλείπονται αυτονόητες παρενθέσεις.

Παρατήρηση: Μια γλώσσα μπορεί να έχει πολλές κανονικές εκφράσεις.

Σχέσεις Κλειστότητας

Κλειστότητα: : Ένα σύνολο γλωσσών A είναι κλειστό ως προς μια δυαδική πράξη $\oplus$ (π.χ. ένωση, τομή, παράθεση) αν και μονο αν για κάθε δυο γλώσσες L_1, L_2 του A, η γλωσσα $L_1 \oplus L_2$ ανήκει επίσης στο A.

Σχέσεις Κλειστότητας

Κλειστότητα: : Ένα σύνολο γλωσσών A είναι κλειστό ως προς μια δυαδική πράξη $\oplus$ (π.χ. ένωση, τομή, παράθεση) αν και μονο αν για κάθε δυο γλώσσες L_1, L_2 του A, η γλωσσα $L_1 \oplus L_2$ ανήκει επίσης στο A.

Ανάλογα ορίζουμε την κλειστότητα για μοναδιαίες πράξεις (π.χ. Kleene star, συμπλήρωμα συνόλου).

Σχέσεις Κλειστότητας

Κλειστότητα: : Ένα σύνολο γλωσσών A είναι κλειστό ως προς μια δυαδική πράξη $\oplus$ (π.χ. ένωση, τομή, παράθεση) αν και μονο αν για κάθε δυο γλώσσες L_1, L_2 του A, η γλωσσα $L_1 \oplus L_2$ ανήκει επίσης στο A.

Ανάλογα ορίζουμε την κλειστότητα για μοναδιαίες πράξεις (π.χ. Kleene star, συμπλήρωμα συνόλου).

Θεωρημα: Οι κανονικές γλώσσες είναι κλειστές ως προς την **ένωση**, **παράθεση**, και **Kleene star**.

Οι κανονικές γλώσσες ενός αλφαβήτου Σ ορίζονται ακριβώς σαν το σύνολο των γλωσσών που περιέχει τις στοιχειώδεις γλωσσες $\emptyset$, και $\{\sigma\}$ για καθε σύμβολο $\sigma \in \Sigma$ και κλείνεται απο την ένωση, παράθεση, και Kleene star.

Σχέσεις Κλειστότητας

Κλειστότητα: : Ένα σύνολο γλωσσών A είναι κλειστό ως προς μια δυαδική πράξη $\oplus$ (π.χ. ένωση, τομή, παράθεση) αν και μονο αν για κάθε δυο γλώσσες L_1, L_2 του A, η γλωσσα $L_1 \oplus L_2$ ανήκει επίσης στο A.

Ανάλογα ορίζουμε την κλειστότητα για μοναδιαίες πράξεις (π.χ. Kleene star, συμπλήρωμα συνόλου).

Θεωρημα: Οι κανονικές γλώσσες είναι κλειστές ως προς την **ένωση**, **παράθεση**, και **Kleene star**.

Οι κανονικές γλώσσες ενός αλφαβήτου Σ ορίζονται ακριβώς σαν το σύνολο των γλωσσών που περιέχει τις στοιχειώδεις γλωσσες $\emptyset$, και $\{\sigma\}$ για καθε σύμβολο $\sigma \in \Sigma$ και κλείνεται απο την ένωση, παράθεση, και Kleene star.

Ερώτηση: Οι κανονικές γλώσσες είναι κλειστές ως προς την τομή; Το συμπλήρωμα;

Γραμματικές και Αυτόματα

Οι κανονικές εκφράσεις αποτελούν ένα είδος γραμματικής: Γενικά γραμματική μιας γλώσσας είναι ένα σύστημα που περιγράφει πως μπορούμε να **παράγουμε** τις συμβολοσειρές μιας γλώσσας.

Γραμματικές και Αυτόματα

Οι κανονικές εκφράσεις αποτελούν ένα είδος γραμματικής: Γενικά γραμματική μιας γλώσσας είναι ένα σύστημα που περιγράφει πως μπορούμε να **παράγουμε** τις συμβολοσειρές μιας γλώσσας.

Ένας άλλος τρόπος για να περιγράψουμε μια γλώσσα είναι τα αυτόματα: Γενικά, αυτόματο μιας γλώσσας L είναι ενας μηχανισμός (αλγόριθμος) που μας επιτρέπει να **αναγνωρίζουμε** με συστηματικό τρόπο αν μια συμβολοσειρά ανήκει στην L ή οχι.

Γραμματικές και Αυτόματα

Οι κανονικές εκφράσεις αποτελούν ένα είδος γραμματικής: Γενικά γραμματική μιας γλώσσας είναι ένα σύστημα που περιγράφει πως μπορούμε να **παράγουμε** τις συμβολοσειρές μιας γλώσσας.

Ένας άλλος τρόπος για να περιγράψουμε μια γλώσσα είναι τα αυτόματα: Γενικά, αυτόματο μιας γλώσσας L είναι ενας μηχανισμός (αλγόριθμος) που μας επιτρέπει να **αναγνωρίζουμε** με συστηματικό τρόπο αν μια συμβολοσειρά ανήκει στην L ή οχι.

Συνήθως, θα μελετάμε κάποιο είδος γραμματικών σε συνδυασμό με το είδος αυτομάτων που αναγνωρίζουν τις αντίστοιχες γλώσσες.