

Συμβολοσειρές (Strings)

Διάλεξη 5: Θεωρία Υπολογισμού · ΗΥ-280

Χαράλαμπος Ε. Τσουρακάκης

Ανακοινώσεις & επόμενες εβδομάδες

Εργασίες

- **Πρόοδος: 24/10**
- **Παράδοση εργασιών: 11/12**
- **Προφορική εξέταση:** την τελευταία εβδομάδα του εξαμήνου.
- Η ακριβής ημερομηνία θα ανακοινωθεί αφού συνεννοηθώ με τη Γραμματεία για τη διαθεσιμότητα χώρων (ελπίζω όχι Σάββατο πάλι!).

Τι ακολουθεί στο μάθημα

- **Αυτή και την επόμενη εβδομάδα:** συμβολοσειρές, γλώσσες, κανονικές γλώσσες και αυτόματα.
- **Την εβδομάδα πριν την πρόοδο:** εισαγωγή στη Lean και θα αφιερώσουμε αρκετό χρόνο κυρίως σε **ασκήσεις και επανάληψη**. 1/67

Συμβολοσειρές (Strings)

Γιατί συμβολοσειρές;

- Κάθε αλγόριθμος και κάθε υπολογιστικό μοντέλο χειρίζεται *ακολουθίες*: πίνακες, λίστες, κείμενο, περιπάτους σε γραφήματα, ακολουθίες εντολών.
- Η είσοδος και η έξοδος κάθε αλγορίθμου είναι, τελικά, μια πεπερασμένη **συμβολοσειρά** (string): τα περιεχόμενα ενός τμήματος της μνήμης.
- Άρα: ο συλλογισμός για τον υπολογισμό απαιτεί συλλογισμό για συμβολοσειρές.

Αλφάβητο και συμβολοσειρά

Ορισμός (Αλφάβητο)

Ένα **αλφάβητο** (alphabet) είναι ένα πεπερασμένο, μη κενό σύνολο συμβόλων Σ . Τα στοιχεία του λέγονται **σύμβολα** (symbols) ή **χαρακτήρες** (characters).

Παραδείγματα:

- Δυαδικό αλφάβητο: $\Sigma = \{0, 1\}$
- Δεκαδικά ψηφία: $\Sigma = \{0, 1, \dots, 9\}$.
- DNA: $\Sigma = \{A, C, G, T\}$.
- Ελληνικοί αλφαβητικοί χαρακτήρες:
 $\Sigma = \{A, B, \dots, \Omega, \alpha, \beta, \dots, \omega, \varsigma\}, \quad |\Sigma| = 49.$
- Το αλφάβητο standard ASCII:
 $\Sigma_{\text{ASCII}} = \{\text{οι 128 χαρακτήρες του ASCII}\}, \quad |\Sigma_{\text{ASCII}}| = 128.$

Ορισμός (Συμβολοσειρά)

Μια **συμβολοσειρά** (string) είναι πεπερασμένη ακολουθία συμβόλων.

- Η άδεια συμβολοσειρά ε δεν περιέχει κανέναν χαρακτήρα.

1. $\Sigma = \{0, 1\}, \quad w = 001011, \quad |w| = 6$

2. $\varepsilon =, |\varepsilon| = 0.$

Python

Στην Python θα γράφαμε $\varepsilon = ""$, ως η άδεια συμβολοσειρά. Αν τρέξουμε `len( $\varepsilon$ )` παίρνουμε το αποτέλεσμα 0.

Ορισμός της συμβολοσειράς αναδρομικά

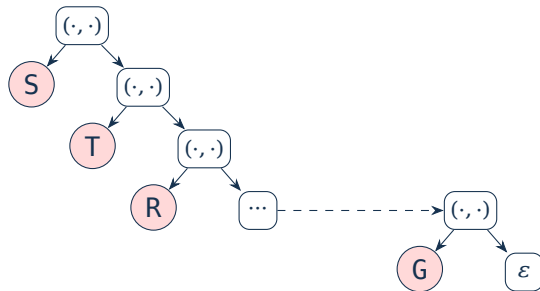
Ορισμός (Συμβολοσειρά αναδρομικά)

Αναδρομικά μια **συμβολοσειρά** (string) ή **λέξη** (word) w πάνω στο Σ ορίζεται ως

- είτε η **κενή συμβολοσειρά** (empty string), που συμβολίζεται ε ,
- είτε ένα διατεταγμένο ζεύγος (a, x) , όπου $a \in \Sigma$ είναι σύμβολο και x είναι συμβολοσειρά πάνω στο Σ .
- Γράφουμε $a \cdot x$ ή απλώς ax αντί για (a, x) .
- Η **STRING** είναι συντομογραφία της τυπικής έκφρασης $(S, (T, (R, (I, (N, (G, \varepsilon))))))$.

Παράδειγμα: η δομή της **STRING**

Η **STRING** είναι συντομογραφία της τυπικής έκφρασης
(**S**, (**T**, (**R**, (**I**, (**N**, (**G**, ϵ))))):



Μήκος συμβολοσειράς

Έστω μια συμβολοσειρά

$$w = w_1 w_2 \cdots w_n, \quad w_i \in \Sigma.$$

Ορισμός (Μήκος συμβολοσειράς)

Το **μήκος** (length) της w , που συμβολίζεται με $|w|$, είναι το πλήθος των συμβόλων της: $|w| = n$.

Μήκος συμβολοσειράς

Έστω μια συμβολοσειρά

$$w = w_1 w_2 \cdots w_n, \quad w_i \in \Sigma.$$

Ορισμός (Μήκος συμβολοσειράς)

Το **μήκος** (length) της w , που συμβολίζεται με $|w|$, είναι το πλήθος των συμβόλων της: $|w| = n$.

Για παράδειγμα, $|01001| = 5$, $|\text{automata}| = 8$.

Υπάρχει επίσης μία μοναδική συμβολοσειρά χωρίς κανένα σύμβολο: η **κενή συμβολοσειρά** ε .

Εξ ορισμού,

$$|\varepsilon| = 0.$$

Προσοχή

Η ε είναι **συμβολοσειρά** μήκους 0. Δεν είναι το κενό σύνολο $\emptyset$.

Η συμβολοσειρά ως συνάρτηση

Μια συμβολοσειρά w μήκους $n = |w|$ πάνω στο αλφάβητο Σ μπορεί ισοδύναμα να θεωρηθεί ως συνάρτηση

$$w : \{1, 2, \dots, n\} \longrightarrow \Sigma.$$

Η τιμή $w(j)$ είναι το **σύμβολο στη θέση j** της συμβολοσειράς.

Ορισμός (Εμφάνιση συμβόλου)

Λέμε ότι το σύμβολο $\sigma \in \Sigma$ **εμφανίζεται στη θέση j** της συμβολοσειράς w αν

$$w(j) = \sigma.$$

Παράδειγμα. Για $w = \text{ABACA}$ έχουμε $|w| = 5$ και

$$w(1) = \text{A}, \quad w(2) = \text{B}, \quad w(3) = \text{A}, \quad w(4) = \text{C}, \quad w(5) = \text{A}.$$

Άρα, για παράδειγμα, το **A** εμφανίζεται στις θέσεις 1, 3, 5.

Ορισμός (Το i -οστό σύμβολο)

Για συμβολοσειρά w και ακέραιο $1 \leq i \leq |w|$, το w_i είναι το i -οστό σύμβολο του w από αριστερά:

$$w_i := \begin{cases} a & \text{αν } w = ax \text{ και } i = 1, \\ x_{i-1} & \text{αν } w = ax \text{ και } i > 1. \end{cases}$$

Ο ακέραιος i λέγεται **δείκτης** (index) του w_i .

- Το w_i είναι μη ορισμένο αν $i < 1$ ή $w = \varepsilon$, άρα (επαγωγικά) και αν $i > |w|$.
- Γράφουμε $w = w_1 w_2 \cdots w_{|w|}$ — αλλά το w μπορεί να έχει ένα σύμβολο ή και κανένα!

- Στον κώδικα, οι δείκτες γράφονται με αγκύλες: $w[i]$.
- Οι αγκύλες καθιερώθηκαν πριν από πάνω από εκατό χρόνια, επειδή δείκτες και εκθέτες ήταν δύσκολο να πληκτρολογηθούν (στη γραφομηχανή!).
- Γράφουμε συμβολοσειρές ως ρητούς πίνακες $w[1 .. n]$, εννοώντας $n = |w|$.

Ορισμός (Υποσυμβολοσειρά)

Μια **υποσυμβολοσειρά** (substring) της w προκύπτει από την w διαγράφοντας μηδέν ή περισσότερα σύμβολα από την αρχή και από το τέλος. Τυπικά: η y είναι υποσυμβολοσειρά της w αν και μόνο αν υπάρχουν συμβολοσειρές x, z με $w = x y z$.

Ορισμός (Η υποσυμβολοσειρά $w[i..j]$)

Η υποσυμβολοσειρά που αρχίζει στο w_i και τελειώνει στο w_j :

$$w[i..j] := \begin{cases} \varepsilon & \text{αν } j < i, \\ w_i \cdot w[i+1..j] & \text{αλλιώς.} \end{cases}$$

Γνήσιες υποσυμβολοσειρές και υπερσυμβολοσειρές

- **Γνήσια** (proper) υποσυμβολοσειρά της w : οποιαδήποτε υποσυμβολοσειρά εκτός από την ίδια την w .
- Αν η y είναι (γνήσια) υποσυμβολοσειρά της w , η w λέγεται (γνήσια) **υπερσυμβολοσειρά** (superstring) της y .

Παράδειγμα

Η **LAUGH** είναι γνήσια υποσυμβολοσειρά της **SLAUGHTER**. Ισοδύναμα, η **SLAUGHTER** είναι γνήσια υπερσυμβολοσειρά της **LAUGH**.

Ορισμός (Πρόθεμα)

Ένα **πρόθεμα** (prefix) της $w[1..n]$ είναι κάθε υποσυμβολοσειρά της μορφής $w[1..j]$. Ισοδύναμα: η p είναι πρόθεμα της w αν και μόνο αν υπάρχει x με $px = w$. **Γνήσιο** πρόθεμα: κάθε πρόθεμα εκτός της w .

Ορισμός (Επίθεμα)

Ένα **επίθεμα** (suffix) της $w[1..n]$ είναι κάθε υποσυμβολοσειρά της μορφής $w[i..n]$. Ισοδύναμα: η s είναι επίθεμα της w αν και μόνο αν υπάρχει x με $xs = w$. **Γνήσιο** επίθεμα: κάθε επίθεμα εκτός της w .

Προθέματα και επιθέματα: παραδείγματα

Παράδειγμα

- **DIE** είναι γνήσιο πρόθεμα της **DIET**.
- **YES** είναι γνήσιο επίθεμα της **EYES**.
- **HE** είναι και γνήσιο πρόθεμα και γνήσιο επίθεμα της **HEADACHE**.

Όλα τα προθέματα και επιθέματα της **ABA**

- Προθέματα: ε , **A**, **AB**, **ABA** ($n + 1 = 4$ προθέματα)
- Επιθέματα: ε , **A**, **BA**, **ABA** ($n + 1 = 4$ επιθέματα)

Η ε και η ίδια η w είναι πάντα προθέματα και επιθέματα της w .

Ορισμός (Υπακολουθία)

Μια **υπακολουθία** (subsequence) της w προκύπτει διαγράφοντας μηδέν ή περισσότερα σύμβολα από *οπουδήποτε* στην w . Τυπικά, η z είναι υπακολουθία της w αν και μόνο αν

- $z = \varepsilon$, ή
- $w = ax$ για κάποιο σύμβολο a και συμβολοσειρά x , και η z είναι υπακολουθία της x , ή
- $w = ax$ και $z = ay$ για κάποιο σύμβολο a και συμβολοσειρές x, y , και η y είναι υπακολουθία της x .

Σε κάθε βήμα είτε *παραλείπουμε* το πρώτο σύμβολο της w , είτε το *ταιριάζουμε* με το πρώτο σύμβολο της z .

Υπακολουθίες: ορολογία και παράδειγμα

- **Γνήσια** υπακολουθία της w : κάθε υπακολουθία εκτός της ίδιας της w .
- Αν η z είναι (γνήσια) υπακολουθία της w , η w λέγεται (γνήσια) **υπερακολουθία** (supersequence) της z .

Παράδειγμα

Όλες οι υπακολουθίες της **ABA**: ε , **A**, **B**, **AA**, **AB**, **BA**, **ABA**.

Η **AA** είναι υπακολουθία, αλλά όχι υποσυμβολοσειρά, της **ABA**.

Υποσυμβολοσειρά $\neq$ υπακολουθία

Μην τις μπερδεύετε!

Κάθε υποσυμβολοσειρά της w είναι και υπακολουθία της w , αλλά όχι κάθε υπακολουθία είναι υποσυμβολοσειρά.

Παράδειγμα

Η **METAL** είναι υπακολουθία, αλλά όχι υποσυμβολοσειρά, της **MEATBALL**:

M E A T B A L L

Τα σύμβολα εμφανίζονται με τη σωστή σειρά, αλλά όχι *συνεχόμενα*.

Για έμφαση, λέμε μερικές φορές πλεοναστικά **συνεχόμενη υποσυμβολοσειρά** (contiguous substring).

Συμβολοσειρές συγκεκριμένου μήκους: Σ^n

Έστω ένα αλφάβητο Σ .

Ορισμός (Σ^n)

Για κάθε $n \in \mathbb{N}$, συμβολίζουμε με Σ^n το σύνολο όλων των συμβολοσειρών πάνω στο Σ που έχουν μήκος ακριβώς n .

Δηλαδή,

$$\Sigma^n = \{w : w \text{ είναι συμβολοσειρά πάνω στο } \Sigma \text{ και } |w| = n\}.$$

Συμβολοσειρές συγκεκριμένου μήκους: Σ^n

Έστω ένα αλφάβητο Σ .

Ορισμός (Σ^n)

Για κάθε $n \in \mathbb{N}$, συμβολίζουμε με Σ^n το σύνολο όλων των συμβολοσειρών πάνω στο Σ που έχουν μήκος ακριβώς n .

Δηλαδή,

$$\Sigma^n = \{w : w \text{ είναι συμβολοσειρά πάνω στο } \Sigma \text{ και } |w| = n\}.$$

Για παράδειγμα, αν $\Sigma = \{0, 1\}$, τότε

- $\Sigma^1 = \{0, 1\}$,
- $\Sigma^2 = \{00, 01, 10, 11\}$,
- $\Sigma^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$.

Τι είναι το Σ^0 ;

Από τον ορισμό,

$$\Sigma^0 = \{w : |w| = 0\}.$$

Αλλά υπάρχει ακριβώς μία συμβολοσειρά μήκους 0: η κενή συμβολοσειρά ε .

Επομένως,

$$\Sigma^0 = \{\varepsilon\}.$$

Τι είναι το Σ^0 ;

Από τον ορισμό,

$$\Sigma^0 = \{w : |w| = 0\}.$$

Αλλά υπάρχει ακριβώς μία συμβολοσειρά μήκους 0: η κενή συμβολοσειρά ε .

Επομένως,

$$\Sigma^0 = \{\varepsilon\}.$$

Μία σημαντική διάκριση

$$\varepsilon \neq \{\varepsilon\} \neq \emptyset.$$

- ε : μία **συμβολοσειρά**,
- $\{\varepsilon\}$: ένα **σύνολο** που περιέχει μία συμβολοσειρά,
- $\emptyset$: ένα **σύνολο** που δεν περιέχει καμία συμβολοσειρά.

Πόσες συμβολοσειρές έχει το Σ^n ;

Έστω

$$|\Sigma| = k.$$

Μια συμβολοσειρά της Σ^n έχει ακριβώς n θέσεις:

$$\square \square \dots \square.$$

Σε κάθε θέση μπορούμε να επιλέξουμε ανεξάρτητα ένα από τα k σύμβολα του Σ .

Άρα, από τον κανόνα του γινομένου,

$$\underbrace{k \cdot k \dots k}_{n \text{ φορές}} = k^n.$$

Επομένως, $|\Sigma^n| = |\Sigma|^n = k^n.$

Πόσες συμβολοσειρές έχει το Σ^n ;

Έστω

$$|\Sigma| = k.$$

Μια συμβολοσειρά της Σ^n έχει ακριβώς n θέσεις:

$$\square \square \dots \square.$$

Σε κάθε θέση μπορούμε να επιλέξουμε ανεξάρτητα ένα από τα k σύμβολα του Σ .

Άρα, από τον κανόνα του γινομένου,

$$\underbrace{k \cdot k \dots k}_{n \text{ φορές}} = k^n.$$

Επομένως, $|\Sigma^n| = |\Sigma|^n = k^n$.

Παρατηρείστε πως αυτό ισχύει και για $n = 0$: $|\Sigma^0| = |\{\varepsilon\}| = 1 = k^0$.

Από τα Σ^n στο Σ^*

Έχουμε ήδη τα σύνολα

$$\Sigma^0, \Sigma^1, \Sigma^2, \Sigma^3, \dots$$

όπου το Σ^n περιέχει όλες τις συμβολοσειρές μήκους n .

Ορισμός (Kleene star του αλφαβήτου)

Το σύνολο **όλων των πεπερασμένων συμβολοσειρών** πάνω στο Σ συμβολίζεται

$$\Sigma^*$$

και ορίζεται ως

$$\Sigma^* = \bigcup_{n=0}^{\infty} \Sigma^n.$$

Δηλαδή,

$$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \dots.$$

Από τα Σ^n στο Σ^*

Έχουμε ήδη τα σύνολα

$$\Sigma^0, \Sigma^1, \Sigma^2, \Sigma^3, \dots$$

όπου το Σ^n περιέχει όλες τις συμβολοσειρές μήκους n .

Ορισμός (Kleene star του αλφαβήτου)

Το σύνολο **όλων των πεπερασμένων συμβολοσειρών** πάνω στο Σ συμβολίζεται

$$\Sigma^*$$

και ορίζεται ως

$$\Sigma^* = \bigcup_{n=0}^{\infty} \Sigma^n.$$

Δηλαδή,

$$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \dots.$$

Για $\Sigma = \{0, 1\}$: $\Sigma^* = \{\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, \dots\}$.

Το Σ^* περιέχει **όλες** τις πεπερασμένες συμβολοσειρές, συμπεριλαμβανομένης της ε :

$$\Sigma^* = \bigcup_{n \geq 0} \Sigma^n.$$

Αν θέλουμε μόνο τις **μη κενές** συμβολοσειρές, γράφουμε

$$\Sigma^+ = \bigcup_{n \geq 1} \Sigma^n.$$

Άρα

$$\Sigma^+ = \Sigma^* \setminus \{\varepsilon\}.$$

Το Σ^* περιέχει **όλες** τις πεπερασμένες συμβολοσειρές, συμπεριλαμβανομένης της ε :

$$\Sigma^* = \bigcup_{n \geq 0} \Sigma^n.$$

Αν θέλουμε μόνο τις **μη κενές** συμβολοσειρές, γράφουμε

$$\Sigma^+ = \bigcup_{n \geq 1} \Sigma^n.$$

Άρα

$$\Sigma^+ = \Sigma^* \setminus \{\varepsilon\}.$$

Για παράδειγμα, αν $\Sigma = \{a, b\}$:

$$\varepsilon \in \Sigma^*, \quad \varepsilon \notin \Sigma^+,$$

ενώ

$$a, ab, bba, aababb \in \Sigma^+.$$

Μικρή άσκηση κατανόησης

Έστω $\Sigma = \{a, b, c\}$. Ποια από τα παρακάτω είναι σωστά;

1. $ab \in \Sigma^2$
2. $ab \in \Sigma^3$
3. $\varepsilon \in \Sigma^0$
4. $\varepsilon \in \Sigma^*$
5. $\varepsilon \in \Sigma^+$
6. $\Sigma \subseteq \Sigma^*$
7. $\Sigma^2 \subseteq \Sigma^*$

Μικρή άσκηση κατανόησης

Έστω $\Sigma = \{a, b, c\}$. Ποια από τα παρακάτω είναι σωστά;

1. $ab \in \Sigma^2$
2. $ab \in \Sigma^3$
3. $\varepsilon \in \Sigma^0$
4. $\varepsilon \in \Sigma^*$
5. $\varepsilon \in \Sigma^+$
6. $\Sigma \subseteq \Sigma^*$
7. $\Sigma^2 \subseteq \Sigma^*$

Απάντηση

Σωστά είναι τα

1, 3, 4, 6, 7.

Συνένωση δύο συμβολοσειρών

Έστω

$$x = x_1 x_2 \cdots x_m \quad \text{και} \quad y = y_1 y_2 \cdots y_n.$$

Ορισμός (Συνένωση)

Η **συνένωση** (concatenation) των x και y είναι η συμβολοσειρά που προκύπτει γράφοντας πρώτα όλα τα σύμβολα της x και αμέσως μετά όλα τα σύμβολα της y :

$$xy = x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n.$$

Συνένωση δύο συμβολοσειρών

Έστω

$$x = x_1 x_2 \cdots x_m \quad \text{και} \quad y = y_1 y_2 \cdots y_n.$$

Ορισμός (Συνένωση)

Η **συνένωση** (concatenation) των x και y είναι η συμβολοσειρά που προκύπτει γράφοντας πρώτα όλα τα σύμβολα της x και αμέσως μετά όλα τα σύμβολα της y :

$$xy = x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n.$$

Για παράδειγμα, αν $x = 011$, $y = 10$, τότε

$$xy = 01110, \quad yx = 10011.$$

Η σειρά έχει σημασία

Γενικά, $xy \neq yx$. Η συνένωση δεν είναι αντιμεταθετική.

Συνένωση και μήκος

Αν

$$x = x_1 \cdots x_m, \quad y = y_1 \cdots y_n,$$

τότε η xy αποτελείται από τα m σύμβολα της x και τα n σύμβολα της y .

Επομένως,

$$|xy| = |x| + |y|.$$

Συνένωση και μήκος

Αν

$$x = x_1 \cdots x_m, \quad y = y_1 \cdots y_n,$$

τότε η xy αποτελείται από τα m σύμβολα της x και τα n σύμβολα της y .

Επομένως,

$$|xy| = |x| + |y|.$$

Παράδειγμα

Για $x = 011$, $y = 10$ έχουμε $|x| = 3$, $|y| = 2$, και πράγματι
 $|xy| = |01110| = 5 = 3 + 2$.

Συνένωση και μήκος

Αν

$$x = x_1 \cdots x_m, \quad y = y_1 \cdots y_n,$$

τότε η xy αποτελείται από τα m σύμβολα της x και τα n σύμβολα της y .

Επομένως,

$$|xy| = |x| + |y|.$$

Παράδειγμα

Για $x = 011$, $y = 10$ έχουμε $|x| = 3$, $|y| = 2$, και πράγματι $|xy| = |01110| = 5 = 3 + 2$.

Επίσης,

$$x\varepsilon = \varepsilon x = x,$$

οπότε η ε είναι το **ουδέτερο στοιχείο** της συνένωσης.

Η σύνδεση είναι προσεταιριστική

Για οποιεσδήποτε συμβολοσειρές x, y, z ,

$$(xy)z = x(yz).$$

Η συνένωση είναι προσεταιριστική

Για οποιεσδήποτε συμβολοσειρές x, y, z ,

$$(xy)z = x(yz).$$

Γιατί;

Αν

$$x = x_1 \cdots x_m, \quad y = y_1 \cdots y_n, \quad z = z_1 \cdots z_k,$$

τότε και οι δύο πλευρές είναι απλώς η συμβολοσειρά

$$x_1 \cdots x_m y_1 \cdots y_n z_1 \cdots z_k.$$

Άρα μπορούμε να γράφουμε χωρίς παρενθέσεις xyz χωρίς καμία ασάφεια.

Προσοχή

- Προσεταιριστική: $(xy)z = x(yz)$.
- Όχι αντιμεταθετική: $xy \neq yx$ γενικά.

Αντιστροφή συμβολοσειράς

Έστω μια συμβολοσειρά

$$x = x_1 x_2 \cdots x_n, \quad x_i \in \Sigma.$$

Ορισμός (Αντιστροφή συμβολοσειράς)

Η **αντιστροφή** (reversal) της x , που συμβολίζεται x^R , είναι η συμβολοσειρά που προκύπτει αν διαβάσουμε την x από **δεξιά προς τα αριστερά**:

$$(x_1 x_2 \cdots x_n)^R = x_n x_{n-1} \cdots x_2 x_1.$$

Για την κενή συμβολοσειρά ορίζουμε

$$\varepsilon^R = \varepsilon.$$

Αντιστροφή συμβολοσειράς

Έστω μια συμβολοσειρά

$$x = x_1 x_2 \cdots x_n, \quad x_i \in \Sigma.$$

Ορισμός (Αντιστροφή συμβολοσειράς)

Η **αντιστροφή** (reversal) της x , που συμβολίζεται x^R , είναι η συμβολοσειρά που προκύπτει αν διαβάσουμε την x από **δεξιά προς τα αριστερά**:

$$(x_1 x_2 \cdots x_n)^R = x_n x_{n-1} \cdots x_2 x_1.$$

Για την κενή συμβολοσειρά ορίζουμε

$$\varepsilon^R = \varepsilon.$$

Παραδείγματα

$$(01101)^R = 10110, \quad (\text{automata})^R = \text{atamotua}.$$

Τι συμβαίνει στη συνένωση όταν αντιστρέφουμε;

Έστω δύο συμβολοσειρές x και y .

Για παράδειγμα, αν

$$x = abc, \quad y = 01,$$

τότε

$$xy = abc01.$$

Άρα

$$(xy)^R = 10cba.$$

Τι συμβαίνει στη συνένωση όταν αντιστρέφουμε;

Έστω δύο συμβολοσειρές x και y .

Για παράδειγμα, αν

$$x = abc, \quad y = 01,$$

τότε

$$xy = abc01.$$

Άρα

$$(xy)^R = 10cba.$$

Όμως

$$x^R = cba, \quad y^R = 10,$$

οπότε

$$y^R x^R = 10cba.$$

Τι συμβαίνει στη συνένωση όταν αντιστρέφουμε;

Ιδιότητα

Για οποιεσδήποτε συμβολοσειρές $x, y \in \Sigma^*$,

$$(xy)^R = y^R x^R.$$

Προσοχή στη σειρά!

Γενικά,

$$(xy)^R \neq x^R y^R.$$

Η αντιστροφή **αντιστρέφει και τη σειρά των κομματιών**.

Απόδειξη 1: $(xy)^R = y^R x^R$

Έστω

$$x = x_1 x_2 \cdots x_m, \quad y = y_1 y_2 \cdots y_n.$$

Τότε, από τον ορισμό της συνένωσης,

$$xy = x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n.$$

Απόδειξη 1: $(xy)^R = y^R x^R$

Έστω

$$x = x_1 x_2 \cdots x_m, \quad y = y_1 y_2 \cdots y_n.$$

Τότε, από τον ορισμό της συνένωσης,

$$xy = x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n.$$

Αντιστρέφοντας αυτή τη συμβολοσειρά παίρνουμε

$$\begin{aligned}(xy)^R &= (x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n)^R \\&= y_n y_{n-1} \cdots y_1 x_m x_{m-1} \cdots x_1 \\&= (y_n y_{n-1} \cdots y_1)(x_m x_{m-1} \cdots x_1) \\&\quad \underbrace{\hspace{1.5cm}}_{y^R} \underbrace{\hspace{1.5cm}}_{x^R} \\&= y^R x^R.\end{aligned}$$

Απόδειξη 1: $(xy)^R = y^R x^R$

Παρατηρείστε ότι η παραπάνω απόδειξη καλύπτει και τις περιπτώσεις $x = \varepsilon$ ή $y = \varepsilon$, αφού $\varepsilon^R = \varepsilon$ και η ε είναι ουδέτερο στοιχείο της συνένωσης.

- Επομένως,

$$(xy)^R = y^R x^R.$$

Αναδρομή και συμβολοσειρές

Γιατί αναδρομικοί ορισμοί συναρτήσεων;

Βασική αρχή

Τα αντικείμενα (οι συμβολοσειρές) ορίζονται αναδρομικά, άρα και οι συναρτήσεις πάνω τους ορίζονται **αναδρομικά**, με την ίδια διάκριση περιπτώσεων: $w = \varepsilon$ ή $w = ax$.

Ορισμός (Μήκος)

Το **μήκος** (length) $|w|$ μιας συμβολοσειράς w είναι το πλήθος των συμβόλων της:

$$|w| := \begin{cases} 0 & \text{αν } w = \varepsilon, \\ 1 + |x| & \text{αν } w = ax. \end{cases}$$

Παράδειγμα

$|\text{FIFTEEN}| = 7$, $|\text{SEVEN}| = 5$, $|\text{5}| = 1$.

- Τυπικά, ένα σύμβολο a και η συμβολοσειρά $a \cdot \varepsilon$ μήκους 1 είναι διαφορετικά αντικείμενα. Στην πράξη δεν τα διακρίνουμε.

Ορισμός (Συνένωση)

Η **συνένωση** (concatenation) $x \cdot y$ (ή απλώς xy) δύο συμβολοσειρών είναι η μοναδική συμβολοσειρά με τα σύμβολα του x με τη σειρά και έπειτα τα σύμβολα του y με τη σειρά. Τυπικά:

$$w \cdot z := \begin{cases} z & \text{αν } w = \varepsilon, \\ a \cdot (x \cdot z) & \text{αν } w = ax. \end{cases}$$

Παράδειγμα

NOW • **HERE** = **NOWHERE**, ενώ **HERE** • **NOW** = **HERENOW**.

Η μεγάλη τελεία • συνενώνει δύο συμβολοσειρές· η μικρή τελεία • βάζει ένα σύμβολο μπροστά από μια συμβολοσειρά.

Παράδειγμα υπολογισμού της συνένωσης

Ας υπολογίσουμε $AB \cdot C$ ακολουθώντας μόνο τον ορισμό:

$$AB \cdot C = (A \cdot B) \cdot C$$

$$= A \cdot (B \cdot C)$$

δεύτερη περίπτωση του ορισμού

$$= A \cdot ((B \cdot \varepsilon) \cdot C)$$

$$= A \cdot (B \cdot (\varepsilon \cdot C))$$

δεύτερη περίπτωση

$$= A \cdot (B \cdot C)$$

πρώτη περίπτωση

$$= ABC$$

Η αναδρομή «ξετυλίγει» το αριστερό όρισμα σύμβολο προς σύμβολο· το δεξί όρισμα μένει ανέγγιχτο.

Δομική Επαγωγή σε συμβολοσειρές

Δύο σχεδιαστικές αρχές

1. Η διάκριση περιπτώσεων της απόδειξης αντικατοπτρίζει τη διάκριση περιπτώσεων του αναδρομικού ορισμού.
2. Η επαγωγική υπόθεση βολεύει συνήθως να είναι ισχυρή όπως θα δούμε

Αρχή 1: Οι περιπτώσεις ακολουθούν τον ορισμό

Αν αποδεικνύουμε κάτι για όλες τις συμβολοσειρές, η απόδειξη έχει **δύο περιπτώσεις**, ακριβώς όπως ο ορισμός της συμβολοσειράς:

- είτε $w = \varepsilon$,
- είτε $w = ax$ για κάποιο σύμβολο a και κάποια συμβολοσειρά x .

Αρχή 2: Ισχυρή επαγωγική υπόθεση

Στην περίπτωση $w = ax$ έχουμε $|x| = |w| - 1 < |w|$ εξ ορισμού του $|w|$, άρα η επαγωγική υπόθεση εφαρμόζεται στο x .

Το «καλούπι» (template) της επαγωγής σε συμβολοσειρές

Θέλουμε να δείξουμε ότι κάθε συμβολοσειρά είναι *μπλουμ*, ό,τι κι αν σημαίνει αυτό.

Απόδειξη

Έστω w μια αυθαίρετη συμβολοσειρά. Υποθέτουμε ότι για κάθε συμβολοσειρά x με $|x| < |w|$, η x είναι *μπλουμ*. Υπάρχουν δύο περιπτώσεις.

- Έστω $w = \varepsilon$. Άρα η w είναι *μπλουμ*
- Έστω $w = ax$ για κάποιο σύμβολο a και συμβολοσειρά x . Η επαγωγική υπόθεση συνεπάγεται ότι η x είναι *μπλουμ* Άρα η w είναι *μπλουμ*

Και στις δύο περιπτώσεις, η w είναι *μπλουμ*.



Τα κενά κουτιά κρύβουν λεπτομέρειες που εξαρτώνται από τον ακριβή ορισμό της ιδιότητας *μπλουμ*.

Πώς γράφουμε τέτοιες αποδείξεις

1. Πρώτα γράφουμε *μηχανικά* το καλούπι, αφήνοντας άφθονο χώρο σε κάθε περίπτωση.
2. Έπειτα συμπληρώνουμε το θεώρημα και την επαγωγική υπόθεση.
3. Μόνο τότε αρχίζουμε πραγματικά να σκεφτόμαστε.

«Μα αυτά είναι προφανή!»

Ναι, τα λήμματα που ακολουθούν είναι ολοφάνερα. Δεν είναι όμως αυτά το ζητούμενο. Προσέξτε πώς κάθε απόδειξη ακολουθεί το καλούπι, και πώς κάθε πρόταση προκύπτει *μηχανικά* από προηγούμενες προτάσεις, ορισμούς και τους κανόνες της λογικής και της αριθμητικής.

Αντιστροφή συμβολοσειράς

Έστω $w \in \Sigma^*$. Η **αντιστροφή** (reversal) της w , που συμβολίζεται με w^R , ορίζεται αναδρομικά.

Ορισμός (Αντιστροφή)

1. **Βάση:** $\varepsilon^R = \varepsilon$.

2. **Αναδρομικό βήμα:**

Αν $w = ua$, $u \in \Sigma^*$, $a \in \Sigma$, τότε

$$(ua)^R = a u^R.$$

Αντιστροφή συμβολοσειράς

Έστω $w \in \Sigma^*$. Η **αντιστροφή** (reversal) της w , που συμβολίζεται με w^R , ορίζεται αναδρομικά.

Ορισμός (Αντιστροφή)

1. **Βάση:** $\varepsilon^R = \varepsilon$.

2. **Αναδρομικό βήμα:**

Αν $w = ua$, $u \in \Sigma^*$, $a \in \Sigma$, τότε

$$(ua)^R = a u^R.$$

Ιδέα

Αφαιρούμε το **τελευταίο σύμβολο** a , το τοποθετούμε στην αρχή και αντιστρέφουμε αναδρομικά την υπόλοιπη συμβολοσειρά u .

Παράδειγμα: αντιστροφή της **DOG**

Θέλουμε να υπολογίσουμε

(DOG)^R.

Παράδειγμα: αντιστροφή της **DOG**

Θέλουμε να υπολογίσουμε

(DOG)^R.

Γράφουμε **DOG** = **DOG**.

Από τον αναδρομικό ορισμό, **(DOG)**^R = **G(DO)**^R.

Παράδειγμα: αντιστροφή της DOG

Θέλουμε να υπολογίσουμε

$$(DOG)^R.$$

Γράφουμε $DOG = DOG$.

Από τον αναδρομικό ορισμό, $(DOG)^R = G(DO)^R$.

Τώρα $DO = DO$, άρα

$$(DO)^R = O(D)^R.$$

Παράδειγμα: αντιστροφή της DOG

Θέλουμε να υπολογίσουμε

$$(DOG)^R.$$

Γράφουμε $DOG = DOG$.

Από τον αναδρομικό ορισμό, $(DOG)^R = G(DO)^R$.

Τώρα $DO = DO$, άρα

$$(DO)^R = O(D)^R.$$

Τέλος,

$$(D)^R = D_{\varepsilon}^R = D.$$

Άρα $(DOG)^R = GOD.$

Παράδειγμα: ξεδιπλώνοντας την αναδρομή

Ολόκληρος ο υπολογισμός μπορεί να γραφτεί ως

$$\begin{aligned}(\text{DOG})^R &= \text{G}(\text{DO})^R \\&= \text{GO}(\text{D})^R \\&= \text{GOD}_\varepsilon^R \\&= \text{GOD}_\varepsilon \\&= \boxed{\text{GOD}}.\end{aligned}$$

Παράδειγμα: Ξεδιπλώνοντας την αναδρομή

Ολόκληρος ο υπολογισμός μπορεί να γραφτεί ως

$$\begin{aligned}(\text{DOG})^R &= \text{G}(\text{DO})^R \\&= \text{GO}(\text{D})^R \\&= \text{GOD}_\varepsilon^R \\&= \text{GOD}_\varepsilon \\&= \boxed{\text{GOD}}.\end{aligned}$$

Σε κάθε αναδρομική κλήση το μήκος μειώνεται κατά 1:

$$3 \longrightarrow 2 \longrightarrow 1 \longrightarrow 0.$$

Από τον αναδρομικό ορισμό σε Python

Ο αναδρομικός ορισμός

$$\varepsilon^R = \varepsilon, \quad (ua)^R = a u^R$$

μεταφράζεται σχεδόν κατά λέξη σε Python.

Αναδρομική υλοποίηση

```
def reverse(w):  
    if len(w) == 0:  
        return ""  
    u = w[:-1]  
    a = w[-1]  
    return a + reverse(u)
```

Από τον αναδρομικό ορισμό σε Python

Ο αναδρομικός ορισμός

$$\varepsilon^R = \varepsilon, \quad (ua)^R = a u^R$$

μεταφράζεται σχεδόν κατά λέξη σε Python.

Αναδρομική υλοποίηση

```
def reverse(w):  
    if len(w) == 0:  
        return ""  
    u = w[:-1]  
    a = w[-1]  
    return a + reverse(u)
```

Για παράδειγμα,

```
reverse("DOG")
```

Ο κώδικας ακολουθεί ακριβώς τον ορισμό

Για $w = \text{DOG}$ έχουμε:

```
reverse("DOG") = G + reverse("DO")  
                = GO + reverse("D")  
                = GOD + reverse("")  
                = GOD.
```

Ο κώδικας ακολουθεί ακριβώς τον ορισμό

Για $w = \text{DOG}$ έχουμε:

```
reverse("DOG") = G + reverse("DO")  
                = GO + reverse("D")  
                = GOD + reverse("")  
                = GOD.
```

Αντιστοιχία μαθηματικών και κώδικα

$\varepsilon^R = \varepsilon \quad \longleftrightarrow \quad \text{if } \text{len}(w) == 0$

$w = ua \quad \longleftrightarrow \quad u = w[:-1], \quad a = w[-1]$

$(ua)^R = au^R \quad \longleftrightarrow \quad \text{return } a + \text{reverse}(u)$

Είναι το παραπάνω μια καλή υλοποίηση **Όχι**. Η Python παρέχει πολύ απλούστερο τρόπο:

Python

```
def reverse(w):  
    return w[::-1]
```

Είναι το παραπάνω μια καλή υλοποίηση **Όχι**. Η Python παρέχει πολύ απλούστερο τρόπο:

Python

```
def reverse(w):  
    return w[::-1]
```

Για παράδειγμα,

```
reverse("DOG")
```

→ **GOD.**

Ξανά $(xy)^R = y^R x^R$: Απόδειξη 2

- Θα δούμε μια δεύτερη απόδειξη της ιδιότητας $(xy)^R = y^R x^R$ χρησιμοποιώντας επαγωγή.
- Για κάθε $n \in \mathbb{N}$, έστω

$$P(n) : \text{για κάθε } w \in \Sigma^* \text{ και κάθε } x \in \Sigma^* \text{ με } |x| = n, \quad (wx)^R = x^R w^R.$$

Βάση της επαγωγής

Για $n = 0$ έχουμε $|x| = 0$, άρα

$$x = \varepsilon.$$

Για κάθε $w \in \Sigma^*$,

$$(wx)^R = (w\varepsilon)^R = w^R.$$

Από την άλλη,

$$x^R w^R = \varepsilon^R w^R = \varepsilon w^R = w^R.$$

Άρα

$$(wx)^R = x^R w^R.$$

Επομένως η $P(0)$ ισχύει.

Ισχυρή επαγωγική υπόθεση

Έστω $n \geq 0$.

Υποθέτουμε ότι η πρόταση ισχύει για **όλες τις συμβολοσειρές μήκους το πολύ n** :

$$\forall k \leq n, \quad \forall u \in \Sigma^*, |u| = k, \quad \forall w \in \Sigma^* : \quad (wu)^R = u^R w^R$$

Θέλουμε να αποδείξουμε ότι ισχύει και για κάθε x με

$$|x| = n + 1.$$

Ισχυρή επαγωγική υπόθεση

Έστω $n \geq 0$.

Υποθέτουμε ότι η πρόταση ισχύει για **όλες τις συμβολοσειρές μήκους το πολύ n** :

$$\forall k \leq n, \quad \forall u \in \Sigma^*, |u| = k, \quad \forall w \in \Sigma^* : \quad (wu)^R = u^R w^R$$

Θέλουμε να αποδείξουμε ότι ισχύει και για κάθε x με

$$|x| = n + 1.$$

Εφόσον $x \neq \varepsilon$, μπορούμε να γράψουμε $x = ua$ για κάποια $u \in \Sigma^*$ και $a \in \Sigma$.

Τότε $|u| = n$.

Επαγωγικό βήμα

Έστω $x = ua$, όπου $|u| = n$ και $a \in \Sigma$.

Για οποιαδήποτε $w \in \Sigma^*$,

$$(wx)^R = (w(ua))^R$$

$$= ((wu)a)^R$$

$$= a(wu)^R$$

$$= a u^R w^R$$

$$= (ua)^R w^R$$

$$= x^R w^R.$$

προσεταιριστικότητα

ορισμός της αντιστροφής

επαγωγική υπόθεση

ορισμός της αντιστροφής

Επαγωγικό βήμα

Έστω $x = ua$, όπου $|u| = n$ και $a \in \Sigma$.

Για οποιαδήποτε $w \in \Sigma^*$,

$$(wx)^R = (w(ua))^R$$

$$= ((wu)a)^R$$

$$= a(wu)^R$$

$$= a u^R w^R$$

$$= (ua)^R w^R$$

$$= x^R w^R.$$

προσεταιριστικότητα

ορισμός της αντιστροφής

επαγωγική υπόθεση

ορισμός της αντιστροφής

Άρα η $P(n+1)$ ισχύει.

Η δομή της απόδειξης

1. Κάνουμε επαγωγή στο **μήκος** της x .

2. **Βάση:**

$$|x| = 0 \implies x = \varepsilon.$$

3. **Ισχυρή επαγωγική υπόθεση:**

υποθέτουμε το αποτέλεσμα για όλες τις συμβολοσειρές μήκους

$$0, 1, \dots, n.$$

4. **Βήμα:** αν $|x| = n + 1$, αφαιρούμε το τελευταίο σύμβολο:

$$x = ua, \quad |u| = n.$$

5. Εφαρμόζουμε την επαγωγική υπόθεση στη **μικρότερη συμβολοσειρά** u .

Χρειαζόμασταν πραγματικά ισχυρή επαγωγή;

Όχι!

Στην απόδειξη της περίπτωσης $|x| = n + 1$ χρησιμοποιήσαμε μόνο την περίπτωση

$$|u| = n.$$

Δεν χρησιμοποιήσαμε τις περιπτώσεις

$$0, 1, \dots, n - 1.$$

Χρειαζόμασταν πραγματικά ισχυρή επαγωγή;

Όχι!

Στην απόδειξη της περίπτωσης $|x| = n + 1$ χρησιμοποιήσαμε μόνο την περίπτωση

$$|u| = n.$$

Δεν χρησιμοποιήσαμε τις περιπτώσεις

$$0, 1, \dots, n - 1.$$

Άρα εδώ θα αρκούσε και η **απλή επαγωγή**.

Γιατί τότε χρησιμοποιούμε ισχυρή επαγωγή;

Το παράδειγμα δείχνει ότι στην ισχυρή επαγωγή μπορούμε, όταν το χρειαζόμαστε, να χρησιμοποιήσουμε **οποιαδήποτε μικρότερη περίπτωση**.

Λήμμα: Το τίποτα δεν αλλάζει τίποτα

Λήμμα

Για κάθε συμβολοσειρά w ισχύει $w \cdot \varepsilon = w$.

Απόδειξη (αρχή)

Έστω w αυθαίρετη. Υποθέτουμε $x \cdot \varepsilon = x$ για κάθε x με $|x| < |w|$.

Υπάρχουν δύο περιπτώσεις.

- Έστω $w = \varepsilon$.

$$w \cdot \varepsilon = \varepsilon \cdot \varepsilon$$

$$= \varepsilon$$

$$= w$$

διότι $w = \varepsilon$

εξ ορισμού της συνένωσης

διότι $w = \varepsilon$

Λήμμα: Το τίποτα δεν αλλάζει τίποτα

Απόδειξη (συνέχεια)

- Έστω $w = ax$ για κάποιο σύμβολο a και συμβολοσειρά x .

$$w \cdot \varepsilon = (a \cdot x) \cdot \varepsilon$$

$$= a \cdot (x \cdot \varepsilon)$$

$$= a \cdot x$$

$$= w$$

$$\text{διότι } w = ax$$

εξ ορισμού της συνένωσης

επαγωγική υπόθεση

$$\text{διότι } w = ax$$

Και στις δύο περιπτώσεις, $w \cdot \varepsilon = w$.



Παρατηρήστε: η επαγωγική υπόθεση εφαρμόζεται στο x επειδή $|x| < |w|$.

Λήμμα: Η συνένωση προσθέτει μήκη

Λήμμα

Για όλες τις συμβολοσειρές w και x ισχύει $|w \bullet x| = |w| + |x|$.

Στήσιμο της απόδειξης

Έστω w, x αυθαίρετες. Υποθέτουμε ότι $|y \bullet x| = |y| + |x|$ για κάθε συμβολοσειρά y με $|y| < |w|$.

Προσοχή

Κάνουμε επαγωγή **μόνο ως προς το** w , όχι ως προς το x — όπως στον ορισμό της $\bullet$, όπου μόνο το αριστερό όρισμα «ξετυλίγεται».

Λήμμα: Η συνένωση προσθέτει μήκη

- Έστω $w = \varepsilon$.

$$|w \cdot x| = |\varepsilon \cdot x|$$

$$= |x|$$

$$= |\varepsilon| + |x|$$

$$= |w| + |x|$$

διότι $w = \varepsilon$

εξ ορισμού της $\cdot$

$|\varepsilon| = 0$ εξ ορισμού του $|\cdot|$

διότι $w = \varepsilon$

- Έστω $w = ay$ για κάποιο σύμβολο a και συμβολοσειρά y .

$$|w \cdot x| = |ay \cdot x|$$

$$= |a \cdot (y \cdot x)|$$

$$= 1 + |y \cdot x|$$

$$= 1 + |y| + |x|$$

$$= |ay| + |x| = |w| + |x|$$

διότι $w = ay$

εξ ορισμού της $\cdot$

εξ ορισμού του $|\cdot|$

επαγωγική υπόθεση

εξ ορισμού του $|\cdot|$, $w = ay$

Και στις δύο περιπτώσεις, $|w \cdot x| = |w| + |x|$.

□

Λήμμα: Η συνένωση είναι προσεταιριστική

Λήμμα

Για όλες τις συμβολοσειρές w, x, y ισχύει $(w \cdot x) \cdot y = w \cdot (x \cdot y)$.

Στήσιμο

Έστω w, x, y αυθαίρετες. Υποθέτουμε $(z \cdot x) \cdot y = z \cdot (x \cdot y)$ για κάθε z με $|z| < |w|$. (Πάλι επαγωγή μόνο ως προς το w .)

- Έστω $w = \varepsilon$.

$$(w \cdot x) \cdot y = (\varepsilon \cdot x) \cdot y$$

$$= x \cdot y$$

$$= \varepsilon \cdot (x \cdot y)$$

$$= w \cdot (x \cdot y)$$

διότι $w = \varepsilon$

εξ ορισμού της $\cdot$

εξ ορισμού της $\cdot$

διότι $w = \varepsilon$

Λήμμα: Η συνένωση είναι προσεταιριστική

- Έστω $w = az$ για κάποιο σύμβολο a και συμβολοσειρά z .

$(w \cdot x) \cdot y = (az \cdot x) \cdot y$	διότι $w = az$
$= (a \cdot (z \cdot x)) \cdot y$	εξ ορισμού της $\cdot$
$= a \cdot ((z \cdot x) \cdot y)$	εξ ορισμού της $\cdot$
$= a \cdot (z \cdot (x \cdot y))$	επαγωγική υπόθεση
$= az \cdot (x \cdot y)$	εξ ορισμού της $\cdot$
$= w \cdot (x \cdot y)$	διότι $w = az$

Και στις δύο περιπτώσεις, $(w \cdot x) \cdot y = w \cdot (x \cdot y)$.



Συμπέρασμα

Από εδώ και πέρα, γράφουμε ελεύθερα wxy χωρίς παρενθέσεις.

Μια εναλλακτική αναδρομική δομή

Το παραπάνω καλούπι δεν είναι το μοναδικό. Μπορούμε να βασίσουμε τη διάκριση περιπτώσεων στην εξής παρατήρηση (η εύκολη απόδειξή της αφήνεται ως άσκηση):

Παρατήρηση

Μια συμβολοσειρά $w \in \Sigma^*$ είναι μη κενή αν και μόνο αν

- είτε $w = a$ για κάποιο σύμβολο $a \in \Sigma$,
- είτε $w = x \cdot y$ για κάποιες μη κενές συμβολοσειρές x και y .
- Στη δεύτερη περίπτωση, το Λήμμα 1.2 δίνει $|x| < |w|$ και $|y| < |w|$, άρα η επαγωγική υπόθεση εφαρμόζεται στο x , στο y , ή και στα δύο.
- Τρεις περιπτώσεις πλέον: $w = \varepsilon$, $w = a$, $w = u \cdot v$.

Η συνένωση είναι προσεταιριστική (ξανά): Στήσιμο και πρώτες περιπτώσεις

Στήσιμο

Έστω w, x, y αυθαίρετες. Υποθέτουμε $(z \cdot x') \cdot y' = z \cdot (x' \cdot y')$ για **όλες** τις συμβολοσειρές x', y' και z με $|z| < |w|$. Χρειαζόμαστε ισχυρότερη επαγωγική υπόθεση από πριν!

- Έστω $w = \varepsilon$. Όπως στην προηγούμενη απόδειξη.
- Έστω $w = a$ για κάποιο σύμβολο a .

$$(w \cdot x) \cdot y = (a \cdot x) \cdot y$$

$$= (a \cdot x) \cdot y$$

$$= a \cdot (x \cdot y)$$

$$= a \cdot (x \cdot y)$$

$$= w \cdot (x \cdot y)$$

$$\text{διότι } w = a$$

$$a \cdot z = a \cdot z \text{ εξ ορισμού της } \cdot$$

$$\text{εξ ορισμού της } \cdot$$

$$a \cdot z = a \cdot z \text{ εξ ορισμού της } \cdot$$

$$\text{διότι } w = a$$

Η συνένωση είναι προσεταιριστική (ξανά): Η περίπτωση $w = u \cdot v$

- Έστω $w = u \cdot v$ για κάποιες μη κενές συμβολοσειρές u και v .

$$(w \cdot x) \cdot y = ((u \cdot v) \cdot x) \cdot y$$

$$= (u \cdot (v \cdot x)) \cdot y$$

$$= u \cdot ((v \cdot x) \cdot y)$$

$$= u \cdot (v \cdot (x \cdot y))$$

$$= (u \cdot v) \cdot (x \cdot y)$$

$$= w \cdot (x \cdot y)$$

$$\text{διότι } w = u \cdot v$$

$$\text{επ. υπόθεση, διότι } |u| < |w|$$

$$\text{επ. υπόθεση, διότι } |u| < |w|$$

$$\text{επ. υπόθεση, διότι } |v| < |w|$$

$$\text{επ. υπόθεση, διότι } |u| < |w|$$

$$\text{διότι } w = u \cdot v$$

Και στις τρεις περιπτώσεις, $(w \cdot x) \cdot y = w \cdot (x \cdot y)$. □

Γιατί χρειάστηκε ισχυρότερη υπόθεση;

Εφαρμόζουμε την υπόθεση με διαφορετικά x' , y' κάθε φορά (π.χ. $x' = v \cdot x$), όχι μόνο με τα αρχικά x , y .

Ορισμός (Επανάληψη συμβόλων)

$$\text{stutter}(w) := \begin{cases} \varepsilon & \text{αν } w = \varepsilon, \\ aa \cdot \text{stutter}(x) & \text{αν } w = ax. \end{cases} \quad \text{Π.χ. } \text{stutter}(\text{ABC}) = \text{AABBCC}.$$

Παράδειγμα απόδειξης: $|\text{stutter}(w)| = 2|w|$

Θεώρημα

Για κάθε συμβολοσειρά w ισχύει $|\text{stutter}(w)| = 2|w|$.

Απόδειξη

Έστω w αυθαίρετη. Υποθέτουμε $|\text{stutter}(x)| = 2|x|$ για κάθε x με $|x| < |w|$.

- Έστω $w = \varepsilon$. Τότε $|\text{stutter}(w)| = |\varepsilon| = 0 = 2|w|$, εξ ορισμού.
- Έστω $w = ax$.

$$\begin{aligned} |\text{stutter}(w)| &= |aa \cdot \text{stutter}(x)| \\ &= |aa| + |\text{stutter}(x)| \\ &= 2 + 2|x| \\ &= 2(1 + |x|) = 2|w| \end{aligned}$$

εξ ορισμού της stutter

Λήμμα η συνένωση προσθέτει μήκη

εξ ορισμού του $|\cdot|$, επ. υπόθεση

εξ ορισμού του $|\cdot|$

Και στις δύο περιπτώσεις, $|\text{stutter}(w)| = 2|w|$.



Ένα νέο είδος εξισώσεων

Έστω Σ ένα αλφάβητο με $1 \in \Sigma$ (π.χ. $\Sigma = \{0, 1\}$).

Πρόβλημα

Να βρεθούν όλες οι συμβολοσειρές $x, y \in \Sigma^*$ για τις οποίες

$$x \mathbf{111} = \mathbf{111} y.$$

- Η πράξη είναι η *συνένωση συμβολοσειρών*. Δλδ. εδώ, το **111** είναι συμβολοσειρά μήκους 3, όχι αριθμός.
- “Να λυθεί” σημαίνει να περιγράψουμε *ολόκληρο* το σύνολο των λύσεων και να αποδείξουμε ότι:
 1. κάθε ζεύγος που δίνουμε είναι πράγματι λύση,
 2. δεν υπάρχουν άλλες λύσεις.

Βήμα 1: μήκη και μία κοινή συμβολοσειρά w

Παίρνοντας μήκη και στα δύο μέλη, $|x| + 3 = 3 + |y|$, επομένως $|x| = |y| =: n$.

Βήμα 1: μήκη και μία κοινή συμβολοσειρά w

Παίρνοντας μήκη και στα δύο μέλη, $|x| + 3 = 3 + |y|$, επομένως $|x| = |y| =: n$.

Θέτουμε $w := x \mathbf{111} = \mathbf{111} y$. Τότε $|w| = n + 3$ και:

- η w **αρχίζει** με **111** και **τελειώνει** με **111**,
- η x αποτελείται από τα πρώτα n σύμβολα της w ,
- η y αποτελείται από τα τελευταία n σύμβολα της w .

Βήμα 1: μήκη και μία κοινή συμβολοσειρά w

Παίρνοντας μήκη και στα δύο μέλη, $|x| + 3 = 3 + |y|$, επομένως $|x| = |y| =: n$.

Θέτουμε $w := x \text{ 111 } = \text{ 111 } y$. Τότε $|w| = n + 3$ και:

- η w **αρχίζει** με **111** και **τελειώνει** με **111**,
- η x αποτελείται από τα πρώτα n σύμβολα της w ,
- η y αποτελείται από τα τελευταία n σύμβολα της w .

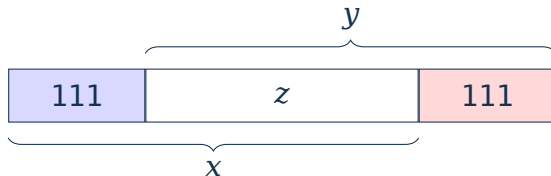
Αντίστροφα, κάθε συμβολοσειρά w μήκους $n + 3$ που αρχίζει και τελειώνει με **111** μάς δίνει με αυτόν τον τρόπο μία λύση.

Άρα: λύσεις $\leftrightarrow$ συμβολοσειρές w που αρχίζουν και τελειώνουν με **111**.

Βήμα 2: η περίπτωση $n \geq 3$ — χωρίς επικάλυψη

Το αρχικό **111** καταλαμβάνει τις θέσεις 1, 2, 3, ενώ το τελικό **111** τις θέσεις $n+1$, $n+2$, $n+3$.

Εφόσον $n+1 \geq 4$, τα δύο αυτά τμήματα είναι **ξένα μεταξύ τους**:



Επομένως $w = \mathbf{111} z \mathbf{111}$, όπου $z \in \Sigma^*$ είναι **αυθαίρετη** και $|z| = n - 3$.

Άρα $x = \mathbf{111} z$, $y = z \mathbf{111}$.

Έλεγχος: $x \mathbf{111} = \mathbf{111} z \mathbf{111} = \mathbf{111} y$. ✓

Βήμα 3: η περίπτωση $n \leq 2$ — επικάλυψη

Τώρα $n + 1 \leq 3$. Άρα το τελικό **111** αρχίζει μέσα στο αρχικό **111** και τα δύο μαζί καλύπτουν όλες τις θέσεις $1, \dots, n + 3$.

Επομένως κάθε σύμβολο της w πρέπει να είναι **1**:

$$w = \mathbf{1}^{n+3}, \quad x = y = \mathbf{1}^n.$$

n	w	x	y
0	111	ε	ε
1	1111	1	1
2	11111	11	11

Για παράδειγμα, όταν $n = 2$, το αρχικό **111** καταλαμβάνει τις θέσεις 1, 2, 3 και το τελικό τις θέσεις 3, 4, 5: **επικαλύπτονται στη θέση 3**.

$$\text{Για } n = 1: \mathbf{1} \cdot \mathbf{111} = \mathbf{1111} = \mathbf{111} \cdot \mathbf{1}. \quad \checkmark$$

Θεώρημα

Ισχύει $x \mathbf{111} = \mathbf{111} y$ αν και μόνο αν ισχύει μία από τις ακόλουθες δύο περιπτώσεις:

1. $x = y = \mathbf{1}^k$, $k \in \{0, 1, 2\}$, ή
2. υπάρχουν $z \in \Sigma^*$ τέτοια ώστε $x = \mathbf{111} z$, $y = z \mathbf{111}$.

Οι δύο οικογένειες είναι ξένες: στην πρώτη $|x| \leq 2$, ενώ στη δεύτερη $|x| \geq 3$.