Μαθηματικά Πληροφορικής

5ο Μάθημα

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστήμιο Αθηνών

Αριθμήσιμα σύνολα

- Πόσα στοιχεία έχει το σύνολο $\{a, b, r, q, x\}$;
- Όσα και το σύνολο $\{1, 2, 3, 4, 5\}$ που είναι υποσύνολο του συνόλου των φυσικών αριθμών $\mathbb{N}$.
- Ένα σύνολο θα το λέμε αριθμήσιμο αν είναι σε 1-1 αντιστοιχία με ένα υποσύνολο των φυσικών αριθμών.

Ορισμός

Ένα σύνολο S λέγεται αριθμήσιμο αν υπάρχει ακολουθία $r_1, r_2, \ldots$ στην οποία να εμφανίζονται όλα τα στοιχεία του S.

Παραδείγματα αριθμήσιμων συνόλων

- Κάθε πεπερασμένο σύνολο είναι αριθμήσιμο.
- Το σύνολο των φυσικών αριθμών $\mathbb{N}$ είναι αριθμήσιμο.

$$1, 2, 3, \ldots$$

- Το σύνολο των ακεραίων είναι αριθμήσιμο.

$$0, 1, -1, 2, -2, 3, -3, \ldots$$

Οι ρητοί αριθμοί

- Είναι αριθμήσιμο το σύνολο των θετικών ρητών $\mathbb{Q}^+$, δηλαδή των κλασμάτων p/q με $p, q \in \mathbb{N}$;
- Ας προσπαθήσουμε να βάλουμε όλα τα κλάσματα σε σειρά ως εξής:
$$\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{2}{1}, \frac{2}{2}, \frac{2}{3}, \ldots, \frac{3}{1}, \ldots$$
- Δείχνει αυτή η σειρά ότι το σύνολο των θετικών ρητών είναι αριθμήσιμο;
- ΟΧΙ! Σε ποιά θέση είναι το $\frac{2}{1}$;

Οι ρητοί αριθμοί (συνέχεια)

- Υπάρχει όμως κατάλληλη ακολουθία.
- Για παράδειγμα:

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{2}{2}, \frac{3}{1}, \frac{1}{4}, \dots$$

- Πρώτα τα κλάσματα με άθροισμα αριθμητή και παρονομαστή ίσο με 2, μετά τα κλάσματα με άθροισμα αριθμητή και παρονομαστή ίσο με 3, κοκ.
- Ένα κλάσμα p/q θα εμφανιστεί όταν απαριθμούμε τα κλάσματα με άθροισμα αριθμητή και παρονομαστή ίσο με $p+q$.
- Επειδή το πλήθος των ζευγαριών των φυσικών αριθμών με άθροισμα το πολύ $p+q$ είναι το πολύ $(p+q)^2$, το κλάσμα p/q, θα εμφανιστεί στα πρώτα $(p+q)^2$ στοιχεία.

Αριθμήσιμα και μη αριθμήσιμα σύνολα Η Μέθοδος της Διαγωνίου Οι πραγματικοί αριθμοί Υπολογισιμότητα

Οι ρητοί αριθμοί (συνέχεια)

- Υπάρχουν πολλές κατάλληλες ακολουθίες.
- Άλλη τέτοια ακολουθία:

$$\frac{1}{1}, \frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{2}{2}, \frac{2}{2}, \frac{1}{3}, \frac{3}{1}, \frac{2}{3}, \frac{3}{2}, \frac{3}{3}, \frac{3}{3}, \frac{1}{4}, \frac{4}{1} \dots$$

- Κάθε κλάσμα p/q ακολουθείται από το αντίστροφό του q/p.
- Τα κλάσματα στις περιττές θέσεις, έχουν αριθμητή μικρότερο ή ίσο του παρονομαστή.
- Σ' αυτά πρώτα είναι τα όλα τα κλάσματα με παρονομαστή 1, μετά όλα τα κλάσματα με παρονομαστή 2, κοκ.
- Επομένως ένα κλάσμα p/q, με $p \leq q$, θα εμφανιστεί όταν απαριθμούνται τα κλάσματα με παρονομαστές q.
- Ένα κλάσμα p/q, με $p > q$, θα εμφανιστεί αμέσως μετά το q/p.

Αριθμήσιμα και μη αριθμήσιμα σύνολα Η Μέθοδος της Διαγωνίου Οι πραγματικοί αριθμοί Υπολογισιμότητα

Συμβολοσειρές

- Αλφάβητο είναι ένα οποιοδήποτε πεπερασμένο σύνολο.
- Έστω αλφάβητο Σ και Σ^* το σύνολο των συμβολοσειρών του.
- Π.χ, $\Sigma = \{0, 1\}$ και $\Sigma^* = \{\varepsilon, 0, 1, 00, 01, 10, 11, \dots\}$.
- Είναι το σύνολο Σ^* αριθμήσιμο;
- Ναι. Ακολουθία: πρώτα οι συμβολοσειρές με μήκος 0 (δηλαδή η ε), μετά όλες τις συμβολοσειρές με μήκος 1, μετά όλες τις συμβολοσειρές με μήκος 2 κοκ.
- Για το παραπάνω παράδειγμα, η ακολουθία είναι

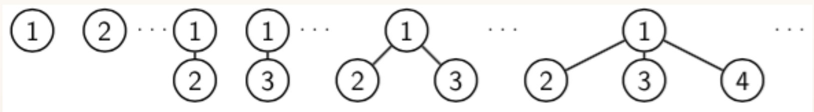
$$\varepsilon, 0, 1, 00, 01, 10, 11, 000, \dots$$

- Επειδή οι συμβολοσειρές με μήκος k είναι πεπερασμένες, μια συμβολοσειρά μήκους k θα εμφανιστεί τελικά στην ακολουθία αυτή.

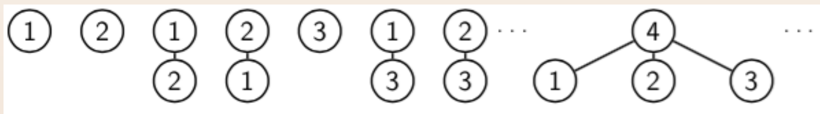
Αριθμήσιμα και μη αριθμήσιμα σύνολα Η Μέθοδος της Διαγωνίου Οι πραγματικοί αριθμοί Υπολογισιμότητα

Δένδρα με ετικέτες (labelled trees)

- Είναι το σύνολο των δένδρων με ετικέτες από το $\mathbb{N}$ αριθμήσιμο;



- Δείχνει η σειρά αυτή ότι τα δένδρα είναι αριθμήσιμα;
- ΟΧΙ! Σε ποια θέση είναι τα δένδρα με 2 κορυφές;
- Υπάρχει όμως κατάλληλη ακολουθία: Πρώτα όλα τα δένδρα με σύνολο κορυφών το $\{1\}$, μετά όλα τα δένδρα με κορυφές στο $\{1, 2\}$, μετά όλα τα δένδρα με κορυφές στο $\{1, 2, 3\}$, κοκ.



Μη αριθμήσιμα σύνολα;

- Είναι όλα τα σύνολα αριθμήσιμα;
- Όχι. Θα δούμε ότι το σύνολο των πραγματικών αριθμών δεν είναι αριθμήσιμο.
- Αλλά πώς δείχνουμε ότι ένα σύνολο δεν είναι αριθμήσιμο;
- Με τη Μέθοδο της Διαγωνίου ή Διαγωνοποίηση.

Η Μέθοδος της Διαγωνίου

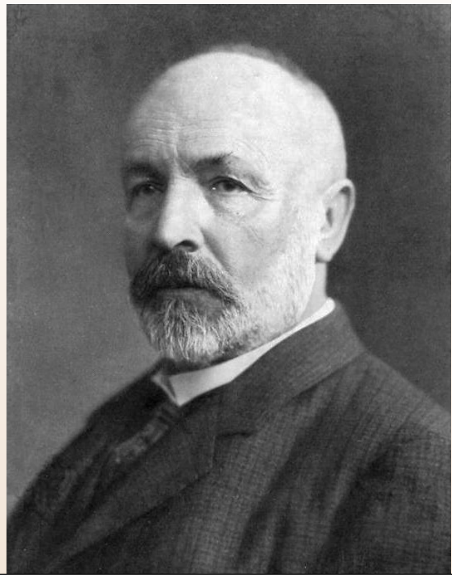
- Δίνονται 5 5-ψήφιοι αριθμοί και μας ζητάνε να βρούμε έναν 5-ψήφιο αριθμό διαφορετικό από αυτούς. Κάθε ψηφίο των 5 αριθμών είναι γραμμένο σε αναποδογυρισμένη κάρτα· θέλουμε να κοιτάξουμε όσο το δυνατόν λιγότερες κάρτες.

$\Rightarrow$

3				
	2			
		9		
			5	
				5

- Ο αριθμός 43066 δεν είναι ένας από τους 5 αριθμούς. Διαφέρει από τον 1ο αριθμό στο 1ο ψηφίο, από τον 2ο στο 2ο ψηφίο και γενικά από τον i-οστό στο i-στο ψηφίο.

Georg Cantor 1845-1918



Το $\mathbb{R}$ δεν είναι αριθμήσιμο

Θεώρημα

Το σύνολο των πραγματικών αριθμών $\mathbb{R}$ δεν είναι αριθμήσιμο, δηλαδή δεν υπάρχει ακολουθία $r_1, r_2, \ldots$ στην όποια να εμφανίζονται όλοι οι πραγματικοί αριθμοί.

Απόδειξη.

- Με εις άτοπο απαγωγή.
- Έστω ότι υπήρχε τέτοια ακολουθία $r_1, r_2, \ldots$
- Ας θεωρήσουμε τη δεκαδική παράσταση των αριθμών αυτών και ας κατασκευάσουμε ένα αριθμό που διαφέρει από τον r_1 στο πρώτο δεκαδικό ψηφίο, από τον r_2 στο δεύτερο δεκαδικό ψηφίο, κοκ.

$\square$

Το $\mathbb{R}$ δεν είναι αριθμήσιμο (συνέχεια)

Συνέχεια.

- Πιο συγκεκριμένα κατασκευάζουμε τον αριθμό a με δεκαδική παράσταση $0.a_1a_2a_3\dots$ όπου

$$a_i = \begin{cases} 1 & \text{αν το } i\text{-στο δεκαδικό ψηφίο του } r_i \text{ είναι διάφορο του } 1 \\ 2 & \text{αν το } i\text{-στο δεκαδικό ψηφίο του } r_i \text{ είναι ίσο με } 1 \end{cases}$$

- Για παράδειγμα, αν

$$\begin{aligned} r_1 &= 0.06542\dots \\ r_2 &= 1.11287\dots \\ r_3 &= 3.14159\dots \\ r_4 &= 1.41421\dots \end{aligned}$$

Αριθμήσιμα και μη αριθμήσιμα σύνολα Η Μέθοδος της Διαγωνίου Οι πραγματικοί αριθμοί Υπολογισιμότητα

Το $\mathbb{R}$ δεν είναι αριθμήσιμο (συνέχεια)

Συνέχεια.

- Πιο συγκεκριμένα κατασκευάζουμε τον αριθμό a με δεκαδική παράσταση $0.a_1a_2a_3\dots$ όπου

$$a_i = \begin{cases} 1 & \text{αν το } i\text{-στο δεκαδικό ψηφίο του } r_i \text{ είναι διάφορο του } 1 \\ 2 & \text{αν το } i\text{-στο δεκαδικό ψηφίο του } r_i \text{ είναι ίσο με } 1 \end{cases}$$

- Για παράδειγμα, αν

$$\begin{aligned} r_1 &= 0.06542\dots \\ r_2 &= 1.11287\dots \\ r_3 &= 3.14159\dots \\ r_4 &= 1.41421\dots \end{aligned}$$

τότε $a = 0.1221\dots$.

Αριθμήσιμα και μη αριθμήσιμα σύνολα Η Μέθοδος της Διαγωνίου Οι πραγματικοί αριθμοί Υπολογισιμότητα

Το $\mathbb{R}$ δεν είναι αριθμήσιμο (συνέχεια)

Συνέχεια.

- Ο νέος αριθμός a έχει διαφορετική δεκαδική παράσταση από κάθε αριθμό της ακολουθίας $r_1, r_2, \dots$.
- Είναι όμως διαφορετικός από κάθε r_i;
- Κάποιοι αριθμοί έχουν δύο δεκαδικές παραστάσεις που από κάποιο σημείο και πέρα αποτελούνται από επαναλαμβανόμενα 0 ή από επαναλαμβανόμενα 9. Για παράδειγμα $1.20000\dots = 1.19999\dots$.
- Φροντίσαμε τα νέα ψηφία a_i να μην περιέχουν ούτε 0 ούτε 9. Έτσι ο νέος αριθμός a, έχει και διαφορετική δεκαδική παράσταση αλλά και διαφορετική τιμή από κάθε r_i.
- Συμπέρασμα: Ο a είναι πραγματικός αριθμός αλλά δεν ανήκει στην ακολουθία $r_1, r_2, \dots$. Άτοπο. Άρα οι πραγματικοί δεν είναι αριθμήσιμοι.

Αριθμήσιμα και μη αριθμήσιμα σύνολα Η Μέθοδος της Διαγωνίου Οι πραγματικοί αριθμοί Υπολογισιμότητα

Θεωρία Υπολογισμού

- Θα δείξουμε ότι δεν υπάρχει πρόγραμμα Pascal ή C ή οποιασδήποτε γενικής γλώσσας προγραμματισμού που να μπορεί να αναλύει τον πηγαίο κώδικα προγραμμάτων της γλώσσας και να απαντά αν τερματίζει ή όχι όταν εκτελεστεί.
- *Ο πηγαίος κώδικας ενός προγράμματος για μας εδώ δεν θα είναι παρά μια συμβολοσειρά (δηλαδή ένα κείμενο)* κάποιου αλφάβητου Σ.
- Θα θεωρήσουμε προγράμματα που παίρνουν για είσοδο μια συμβολοσειρά του Σ.
- Έστω $P_1, P_2, \dots$ τα προγράμματα αυτά. Τα προγράμματα είναι συμβολοσειρές, άρα είναι αριθμήσιμα.

Alan Turing 1912-1954

Θεωρία Υπολογισμού

- Ένα τέτοιο πρόγραμμα μπορεί να τερματίζει ή όχι. Τί θα κάνουν τα προγράμματα αν τους δώσουμε για είσοδο τους πηγαίους κώδικες;
- Ας ορίσουμε την συνάρτηση

$$g(P_i, P_j) = \begin{cases} \text{true} & \text{αν το πρόγραμμα } P_i \text{ τερματίζει με είσοδο} \\ & \text{τη συμβολοσειρά } P_j \\ \text{false} & \text{σε κάθε άλλη περίπτωση} \end{cases}$$

Θεωρία Υπολογισμού

	P_1	P_2	P_3	P_4	P_5
P_1	$g(P_1, P_1)$	$g(P_1, P_2)$	$g(P_1, P_3)$	$g(P_1, P_4)$	
P_2	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
P_3	$\cdots$	$\cdots$	$\cdots$	$g(P_3, P_4)$	
P_4	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
P_5	$\cdots$	$\cdots$	$\cdots$	$\cdots$	

- Χρησιμοποιώντας τη Μέθοδο της Διαγωνίου θεωρούμε τη γραμμή που έχει αντίστροφες τιμές από αυτές της διαγώνιου (από true σε false και το αντίστροφο).
- Η γραμμή που κατασκευάζεται με αυτό τον τρόπο, έχει στην i-οστή θέση true αν και μόνο αν $g(P_i, P_i) = \text{false}$.
- Η γραμμή αυτή δεν υπάρχει στον πίνακα γιατί διαφέρει από την 1η γραμμή στην 1η θέση, από τη 2η γραμμή στη 2η θέση κοκ.
- Δηλαδή, δεν υπάρχει πρόγραμμα που να υπολογίζει το "$not\ g(P_i, P_i)$".

Θεωρία Υπολογισμού

- Ισοδύναμα, δεν υπάρχει πρόγραμμα H τέτοιο ώστε για κάθε συμβολοσειρά P_i

  το H με είσοδο P_i επιστρέφει true
  όταν και μόνο όταν
  το πρόγραμμα P_i με είσοδο P_i δεν τερματίζει

- Συμπεραίνουμε ότι δεν είναι υπολογίσιμο αν ένα πρόγραμμα τερματίζει ή όχι, γιατί διαφορετικά θα μπορούσαμε να υπολογίσουμε αν το P_i τερματίζει με είσοδο τον πηγαίο κώδικά του.