ΕΠΛ 211:
Θεωρία Υπολογισμού και Πολυπλοκότητας

Διάλεξη 15: Διαγνωσιμότητα (Επιλυσιμότητα) II

Τι θα κάνουμε σήμερα…

- Επιλύσιμα Προβλήματα σχετικά με Ασυμφραστικές Γλώσσες (4.1.2)
- Το Πρόβλημα του Τερματισμού (4.2)

Προβλήματα σχετικά με Ασυμφραστικές Γλώσσες

- Υπολογιστικά προβλήματα για ασυμφραστικές γραμματικές. Αλγόριθμοι που ελέγχουν εάν:
  - Ασυμφραστική γραμματική παράγει μια λέξη
  - Γλώσσα μιας ασυμφραστικής γραμματικής είναι η κενή
  - Ασυμφραστικές γραμματικές διαγνώσιμες

Παραγωγή/CFG

- ΠΑΡΑΓΩΓΗ/CFG={<G,w> | το G είναι μια CFG που παράγει την λέξη w}
- Ιδέα 1: Διατρέχουμε όλες τις πιθανές παραγωγές της G και να ελέγξουμε εάν η w είναι μια από αυτές
  - Πρόβλημα: Μπορεί να υπάρχει άπειρο πλήθος παραγωγών
  - Εάν δεν παράγουμε την w ο αλγόριθμος δεν τερματίζει
  - Σε αυτή την περίπτωση η TM αναγνωρίζει αλλά δεν διαγιγνώσκει
- Ιδέα 2: Εάν η G είναι σε κανονική μορφή Chomsky τότε οποιαδήποτε παραγωγή της w αποτελείται από 2n-1 βήματα (όπου n=|w|).
  - Μπορούμε να ελέγξουμε μόνο παραγωγές 2n-1 βημάτων που είναι πεπερασμένες σε αριθμό

Παραγωγή/CFG

- S=‘ Για είσοδο <G,w>, όπου G μια CFG και w μια λέξη:
  1. Μετατρέπουμε την G σε μια ισοδύναμη γραμματική σε κανονική μορφή Chomsky.
  2. Συντάσσουμε όλες τις παραγωγές 2n-1 βημάτων
  3. Εάν κάποια από τις παραγωγές είναι η w, αποδεχόμαστε, αλλιώς απορρίπτουμε.’

- Η γλώσσα ΠΑΡΑΓΩΓΗ/CFG είναι **διαγνώσιμη (αναδρομική)**

Κενότητα/CFG

- ΚΕΝΟΤΗΤΑ/CFG={<G> | η G είναι μια CFG και L(G)=∅}

- Ιδέα 1: Να ελέγξουμε όλες τις πιθανές λέξεις w
  - Άπειρες λέξεις και επομένως μπορεί να μην τερματίσουμε

- Ιδέα 2: Πότε μια CFG παράγει κάποια από τις εισόδους του;
  - Όταν από την εναρκτήρια μεταβλητή μπορούμε να παραγάγουμε λέξεις τερματικών συμβόλων
  - Ελέγχουμε αν από κάθε μεταβλητή μπορούμε να παράγουμε λέξεις τερματικών συμβόλων

- R=‘ Για είσοδο <G>, όπου G μια CFG:
  1. Σημαδεύουμε όλα τα τερματικά σύμβολα της G.
  2. Επαναλαμβάνουμε μέχρις ότου να μην σημαδευτούν όλες οι μεταβλητές:
  3. Σημαδεύουμε κάθε μεταβλητή A->$U_1 \ldots U_k$ όπου $U_1 \ldots U_k$ είναι ήδη σημαδεμένα.
  4. Εάν η εναρκτήρια δεν είναι σημαδεμένη, αποδεχόμαστε, αλλιώς απορρίπτουμε.’

- Η γλώσσα ΚΕΝΟΤΗΤΑ/CFG είναι **διαγνώσιμη (αναδρομική)**

Ισοδυναμία/CFG

- ΙΣΟΔΥΝΑΜΙΑ/CFG={<G,H> | τα G,H είναι CFG και L(G)=L(H)}

- Ιδέα 1: να χρησιμοποιήσουμε παρόμοια τακτική με τον αλγόριθμο που δείχνει την ισοδυναμία των κανονικών γλωσσών
  - Αδύνατον αφού οι ασυμφραστικές γλώσσες δεν είναι κλειστές κάτω από τις πράξεις του συμπληρώματος και της τομής

- Πραγματικότητα: Η γλώσσα ΙΣΟΔΥΝΑΜΙΑ/CFG δεν είναι διαγνώσιμη

Κάθε Ασυμφραστική είναι διαγνώσιμη

- Έστω A μια ασυμφραστική γλώσσα και θέλουμε να δώσουμε αλγόριθμο που να την διαγιγνώσκει

- Αφού A είναι ασυμφραστική τότε
  - Υπάρχει αυτόματο στοίβας που την αναγνωρίζει
  - Υπάρχει ασυμφραστική γραμματική που την παράγει

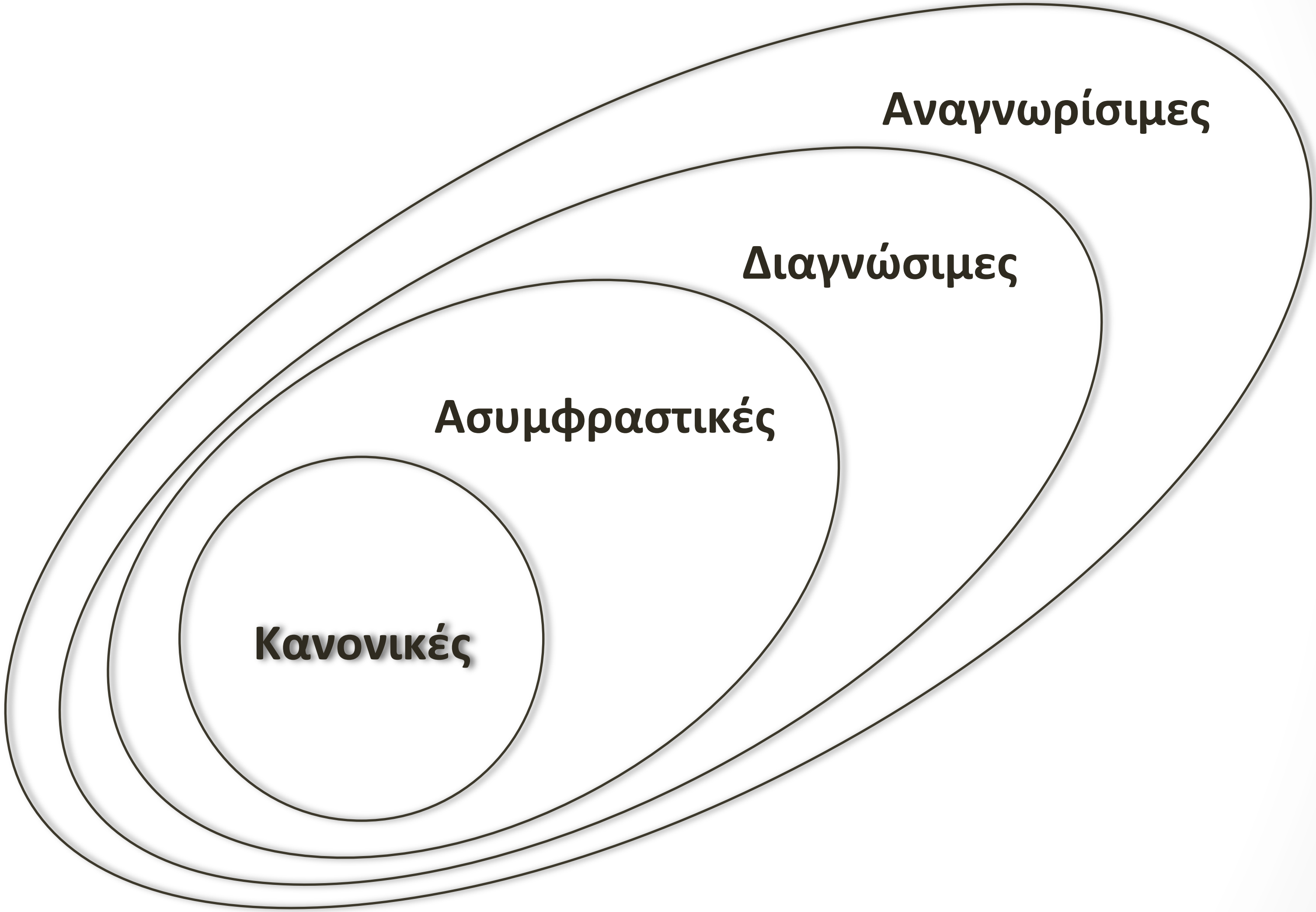
- Ιδέα 1: Να προσομοιωθεί το αυτόματο στοίβας που αναγνωρίζει την A από μια μηχανή Turing
  - Σχετικά απλή προσομοίωση
  - Όμως: Αν κάποιοι κλάδοι του αυτομάτου γράφουν και διαβάζουν από την στοίβα επ’απειρον τότε η TM δεν τερματίζει.

Κάθε Ασυμφραστική είναι διαγνώσιμη

- Ιδέα 2: Να ελέγξουμε αν η γραμματική που παράγει την Α μπορεί να προσομοιωθεί από ΤΜ
  - Χρήση μηχανής που διαγιγνώσκει την γλώσσα ΠΑΡΑΓΩΓΗ/CFG

- M_G=‘ Για είσοδο w:
  1. Εκτελούμε την TM S πάνω στην είσοδο <G,w>
  2. Εάν η S αποδεχθεί τότε αποδεχόμαστε, αλλιώς απορρίπτουμε.’

Σχέση Γλωσσών



Αποδοχή/TM

- ΑΠΟΔΟΧΗ/TM={<M,w> |η M είναι μια TM που αποδέχεται την λέξη w}

Θεώρημα:
Η γλώσσα ΑΠΟΔΟΧΗ/TM είναι μη διαγνώσιμη.

- Γενικές TM, οι «αναγνωριστές», είναι **ισχυρότερες** από του διαγνώστες
  - Απαιτώντας από μια μηχανή να τερματίζει σε κάθε είσοδο (διαγνώστης) περιορίζουμε την κλάση των γλωσσών που μπορεί να αναγνωρίσει.

Αποδοχή/TM - Αναγνωρίσιμη

- Η TM που μπορεί να αναγνωρίσει την ΑΠΟΔΟΧΗ/TM είναι:
- U=‘ Για είσοδο <M,w>, όπου M μια TM και w μια λέξη:
  1. Προσομοιώνουμε την M για είσοδο w.
  2. Εάν η M μεταβεί ποτέ στην κατάσταση αποδοχής, αποδεχόμαστε. Εάν η M μεταβεί ποτέ στην κατάσταση απόρριψης απορρίπτουμε.’
- Αν η M εγκλωβίζεται για είσοδο w τότε και η U εγκλωβίζεται για είσοδο <M,w>

- Η μηχανή U είναι ένα παράδειγμα **Καθολικής TM**
  - Μπορεί να προσομοιώνει οποιαδήποτε άλλη μηχανή με αφετηρία την περιγραφή της

Μέθοδος Διαγωνιοποίησης

- Δείξαμε ότι η ΑΠΟΔΟΧΗ/ΤΜ είναι αναγνωρίσιμη. Πως ξέρουμε όμως ότι δεν είναι διαγνώσιμη;
  - Πρέπει να ελέγξουμε όλες τις πιθανές ΤΜ και να δείξουμε ότι καμιά δεν αποδέχεται όταν η Μ δεν τερματίζει.
  - ΠΡΟΒΛΗΜΑ ΤΟΥ ΤΕΡΜΑΤΙΣΜΟΥ

- Η απόδειξη βασίζεται στη τεχνική της **διαγωνιοποίησης**
  - Επινοήθηκε από τον μαθηματικό Georg Cantor, 1873
  - Προσπαθούσε να συγκρίνει τα μεγέθη 2 άπειρων συνόλων

Μέθοδος Διαγωνιοποίησης

- Πως μπορούμε να ξέρουμε αν δυο πεπερασμένα σύνολα είναι ισομεγέθη;
  - Μετρούμε τα στοιχεία τους
  - Cantor: Αν μπορούμε να συντεριάξουμε κάθε στοιχείο του ενός με ένα στοιχείο του άλλου

Ορισμός Διαγωνιοποίησης:
Δύο σύνολα A και B είναι ισομεγέθη εάν υπάρχει ένα-προς-ένα και επί συνάρτηση (αντιστοιχία) από το A στο B.

Παράδειγμα 1

- Έστω τα σύνολα:
  - N = {1, 2, 3, ...} φυσικοί αριθμοί
  - A = {2, 4, 6,...} άρτιοι αριθμοί
  - Είναι τα δύο σύνολα ισομεγέθη;
- Θεωρήστε την αντιστοιχία f(n)=2n από το N στο A

n	**f(n)**
1	2
2	4
3	6
…	…

- Κάθε στοιχείο του N αντιστοιχεί σε ένα μοναδικό στοιχείο του A
  - Άρα N και A σύμφωνα με τον Cantor είναι ισομεγέθη!!

Αριθμήσιμα και Υπεραριθμήσιμα Σύνολα

Ορισμός:
Ένα σύνολο λέγεται αριθμήσιμο εάν είναι πεπερασμένο ή ισομέγεθες με το N.

Ορισμός:
Ένα σύνολο λέγεται υπεραριθμήσιμο εάν δεν είναι πεπερασμένο ή ισομέγεθες με το N.

Παράδειγμα 2

- Έστω το σύνολο R των πραγματικών αριθμών
- Θεώρημα: Το R είναι υπεραριθμήσιμο
- Έστω ότι υπάρχει αντιστοιχεία f από το N στο R

n	f(n)
1	3.14159
2	55.55555
3	0.12345
…	…

- Θέλουμε να κατασκευάσουμε ένα αριθμό x που δεν αντιστοιχεί σε κανένα από τους αριθμούς στο N
  - Ο x πρέπει να διαφέρει από όλους τους πραγματικούς αριθμούς που έχουν αντιστοιχία

Παράδειγμα 2

- Κατασκευή x:
  - Επιλέγουμε το x μεταξύ [0,1] τ.ω.
  - Για το n-οστό δεκαδικό ψηφίο του x επιλέγουμε ένα αριθμό διαφορετικό από το n-οστό ψηφίο του δεκαδικού που αντιστοιχεί στο n.
  - Π.χ. Το πρώτο δεκαδικό του x είναι 4 που είναι διαφορετικό από το πρώτο δεκαδικό του f(1)=3.<u>1</u>4159

n	f(n)
1	3.<u>1</u>4159
2	55.5<u>5</u>555
3	0.123<u>4</u>5
…	…

- Άρα ο αριθμός x=0.4641… δεν συμπίπτει με τον αριθμό f (n) για οποιοδήποτε n

Μη Αναγνωρίσιμες Γλώσσες

- Πόρισμα: Υπάρχουν μη αναγνωρίσιμες γλώσσες
- Υπεραριθμήσιμο πλήθος γλωσσών
- Αριθμήσιμο πλήθος ΤΜ
  - Κάθε ΤΜ αναγνωρίζει μια μόνο γλώσσα
- Υπάρχουν γλώσσες που δεν αναγνωρίζει καμιά ΤΜ

Απόδειξη: Πρόβλημα του τερματισμού

- Έστω ότι η ΑΠΟΔΟΧΗ/ΤΜ είναι διαγνώσιμη
  - Άρα υπάρχει ΤΜ Η που την διαγιγνώσκει

- Η=‘ Για είσοδο <M,w>, όπου Μ μια ΤΜ και w μια λέξη:
  1. Προσομοιώνουμε την Μ για είσοδο w.
  2. Εάν η Μ μεταβεί ποτέ στην κατάσταση αποδοχής, αποδεχόμαστε. Εάν η Μ δεν αποδέχεται, είτε επειδή έφτασε σε κατάσταση απόρριψης είτε επειδή δεν τερματίζει, απορρίπτουμε.’

Απόδειξη: Πρόβλημα του τερματισμού

- Δημιουργούμε μια μηχανή Δ που χρησιμοποιεί την Η ως υποπρόγραμμα
  - Η Δ ελέγχει πως συμπεριφέρεται μια TM M αν της δώσουμε ως είσοδο την λέξη που αντιστοιχεί στην κωδικοποίησή της <M>.
  - Η Δ απορρίπτει εάν η Μ αποδέχεται την <M>
  - Η Δ αποδέχεται εάν η Μ απορρίπτει την <M>

- Δ=‘ Για είσοδο <M>, όπου Μ μια TM:
  1. Εκτελούμε την Η για είσοδο <M,<M>>.
  2. Εάν η Η απορρίπτει, αποδεχόμαστε αλλιώς εάν η Η αποδέχεται, απορρίπτουμε.’

Απόδειξη: Πρόβλημα του τερματισμού

- Άρα η Δ διαγραμματικά δουλεύει ως εξής:
- $$Δ(<M>) = \begin{cases} \textbf{αποδοχή} & \text{εάν η M δεν αποδέχεται την λέξη } <M> \\ \textbf{απόρριψη} & \text{εάν η M αποδέχεται την λέξη } <M> \end{cases}$$
- Τι γίνεται αν τρέξουμε την Δ με τον εαυτό της;
- $$Δ(<Δ>) = \begin{cases} \textbf{αποδοχή} & \text{εάν η Δ δεν αποδέχεται την λέξη } <Δ> \\ \textbf{απόρριψη} & \text{εάν η Δ αποδέχεται την λέξη } <Δ> \end{cases}$$
- **ΑΤΟΠΟ**

Μια μη αναγνωρίσιμη γλώσσα

- Θυμηθείτε: Μια γλώσσα είναι διαγνώσιμη εάν και μόνο εάν είναι αναγνωρίσιμη (αναδρομικά απαριθμήσιμη) και συμπληρωματικά αναγνωρίσιμη (συναναδρομικά απαριθμήσιμη).

- Το συμπλήρωμα της ΑΠΟΔΟΧΗ/ΤΜ δεν ειναι αναγνωρίσιμη γλώσσα
  - Η ΑΠΟΔΟΧΗ/ΤΜ είναι αναγνωρίσιμη
  - Αν το συμπλήρωμα της ΑΠΟΔΟΧΗ/ΤΜ ήταν αναγνωρίσιμη τότε η ΑΠΟΔΟΧΗ/ΤΜ θα ήταν διαγνώσιμη
  - Αποδείξαμε ότι η ΑΠΟΔΟΧΗ/ΤΜ δεν ειναι διαγνώσιμη

Ερωτήσεις;