Τυχαιότητα (Randomness) I

- Χρησιμοποιώντας το μοντέλο **δένδρων υπολογισμού**, θα ορίσουμε κλάσεις πολυπλοκότητας που βασίζονται στις **πιθανότητες**, με βάση τυχαίες επιλογές.
- Αυτή η προσέγγιση είναι πολύ χρήσιμη από πρακτική άποψη, αφού σε πολλές εφαρμογές, είναι ικανοποιητικός ένας αλγόριθμος ο οποίος κάνοντας κάποιες τυχαίες επιλογές, δίνει στις περισσότερες των περιπτώσεων το σωστό αποτέλεσμα.
- Ένας **πιθανοκρατικός αλγόριθμος** είναι συνήθως πιο απλός στην διατύπωσή του και στην πράξη πιο αποδοτικός από έναν αντίστοιχο ντετερμινιστικό που επιλύει το ίδιο πρόβλημα. Για παράδειγμα, απλοί πιθανοκρατικοί αλγόριθμοι για τον έλεγχο αν ένας αριθμός είναι πρώτος υπάρχουν από την δεκαετία του 1970 και χρησιμοποιούνται στην πράξη έναντι πιο περίπλοκων ντετερμινιστικών τύπου AKS.
- Στα πλαίσια του μοντέλου δένδρων υπολογισμού, θα θεωρήσουμε ότι η επιλογή σε κάθε κόμβο του δένδρου γίνεται τυχαία με πιθανότητα $1/2$ για κάθε παιδί του κόμβου. Για να δείξουμε ότι η *«συντριπτική» πλειοψηφία* των υπολογισμών δίνει το σωστό αποτέλεσμα, εισαγάγουμε έναν νέο ποσοδείκτη, τον $\exists^+$.

Τυχαιότητα (Randomness) II

- Με την βοήθεια του $\exists^+$, ορίζουμε την κλάση BPP, από το **Bounded Probabilistic Polynomial**:

Ορισμός ($\mathrm{BPP} = (\exists^+, \exists^+)$)

$$L \in \mathrm{BPP} \iff \exists R \in \mathrm{P}: \begin{cases} x \in L & \implies \exists^+ y R(x,y) \\ x \notin L & \implies \exists^+ y \neg R(x,y) \end{cases}$$

Τυχαιότητα (Randomness) III

- Με άλλα λόγια, σε ένα δέντρο για την κλάση BPP έχουμε την «συντριπτική» πλειοψηφία των φύλλων να δίνει το σωστό αποτέλεσμα. Στον παραπάνω ορισμό, δεν έχει μεγάλη σημασία ο ακριβής ορισμός της «συντριπτικής» πλειοψηφίας, αλλά πρέπει να είναι οπωσδήποτε *φραγμένος* (εξ ου και το 'bounded' του BPP) πάνω από το $1/2$. Το ποσοστό της πλειοψηφίας μπορεί να είναι, ενδεικτικά, μεγαλύτερο από $1/2 + \varepsilon$, $1/2 + 1/p(|x|)$, $2/3$, 99%, $1 - 2^{-p(|x|)}$ (όπου $p(|x|) > 1$). Αυτή η δυνατότητα επιλογής υπάρχει, επειδή με πολυωνυμικές επαναλήψεις του αντίστοιχου αλγορίθμου, είναι δυνατόν να αυξήσουμε την πιθανότητα επιτυχίας, όσο θέλουμε. Αλγόριθμοι BPP ονομάζονται **Monte Carlo** ή αλλιώς two-sided error, επειδή ανεξάρτητα από το αποτέλεσμα (ναι ή όχι), υπάρχει κάποια πιθανότητα λάθους. Είναι προφανές ότι η κλάση BPP είναι κλειστή ως προς συμπλήρωμα.

Τυχαιότητα (Randomness) IV

- Ας θεωρήσουμε τώρα αλγορίθμους οι οποίοι κάνουν λάθος μόνον για την μία απάντηση (one sided error). Έτσι, προκύπτει η κλάση RP (**Randomized Polynomial**):

Ορισμός ($\mathrm{RP} = (\exists^+, \forall)$)

$$L \in \mathrm{RP} \iff \exists R \in \mathrm{P}: \begin{cases} x \in L & \implies \exists^+ y R(x,y) \\ x \notin L & \implies \forall y \neg R(x,y) \end{cases}$$

Σε αυτήν την κλάση, αν ο αντίστοιχος RP αλγόριθμος δώσει απάντηση «ναι» (δηλαδή το κατηγόρημα R υπολογιστεί αληθές), είμαστε σίγουροι ότι $x \in L$. Αντίθετα, η απάντηση «όχι» του RP αλγορίθμου δεν είναι «σίγουρη».

Προφανώς, ισχύουν: $\mathrm{RP} \subseteq \mathrm{BPP}$, $\mathrm{coRP} \subseteq \mathrm{BPP}$, αλλά δεν γνωρίζουμε αν $\mathrm{RP} = \mathrm{coRP}$.

Τυχαιότητα (Randomness) V

- Μία άλλη πολύ χρήσιμη κλάση, είναι αυτή που ορίζεται με τομή των RP και coRP, η $\mathrm{ZPP} = \mathrm{RP} \cap \mathrm{coRP}$. Η ονομασία προέρχεται από το **Zero error Probabilistic Polynomial**, γιατί μπορεί εύκολα να δειχτεί ότι ένα πρόβλημα είναι στο ZPP αν υπάρχει πιθανοκρατικός αλγόριθμος ο οποίος τρέχει σε αναμενόμενο πολυωνυμικό χρόνο και δίνει πάντοτε σωστή απάντηση. Πράγματι, αν ένα πρόβλημα είναι στο ZPP, σημαίνει ότι έχουμε ένα RP και έναν coRP αλγόριθμο για αυτό, οπότε αρκεί να τρέχουμε εναλλακτικά τους δύο αλγορίθμους, μέχρι ο ένας να δώσει την «σίγουρή» του απάντηση. Βέβαια, μπορεί να χρειαστεί να τρέχουμε εναλλακτικά τους δύο αλγορίθμους για πάντα, αλλά με μεγάλη πιθανότητα θα έχουμε μία «σίγουρη» απάντηση, μετά από μερικές επαναλήψεις. Εναλλακτικά, μπορούμε να πούμε ότι ένας ZPP αλγόριθμος έχει τρεις εξόδους: «ναι», «όχι» (για τις «σίγουρες» απαντήσεις), και «δεν ξέρω» (για τις όχι «σίγουρες»).

  Οι αλγόριθμοι στο ZPP ονομάζονται **Las Vegas**.

Τυχαιότητα (Randomness) VI

- Δεδομένου ότι υπάρχουν αρκετοί πιθανοκρατικοί αλγόριθμοι ευρείας χρήσης για πρακτικά προβλήματα, πολλοί τοποθετούν τους εφικτούς (feasible) υπολογισμούς πάνω από το P, στις πιθανοτικές κλάσεις BPP, RP, ZPP.

  Πάντως, δεν γνωρίζουμε αν υπάρχουν πλήρη προβλήματα για τις κλάσεις που ορίστηκαν παραπάνω (BPP, RP, ZPP).
- Αν τώρα το ποσοστό λάθους ενός πιθανοκρατικού αλγορίθμου δεν φραχθεί μακριά από το $1/2$, τότε έχουμε απλώς την βεβαιότητα ότι στο μοντέλο δένδρων υπολογισμού παραπάνω από τα μισά υπολογιστικά μονοπάτια δίνουν την σωστή απάντηση. Για να δηλώσουμε το παραπάνω χρησιμοποιούμε τον ποσοδείκτη $\exists_{1/2}$. Για unbounded two-sided error, έχουμε την κλάση PP (Probabilistic Polynomial):

Ορισμός ($\mathrm{PP} = (\exists_{1/2}, \exists_{1/2})$)

$$L \in \mathrm{PP} \iff \exists R \in \mathrm{P}: \begin{cases} x \in L & \implies \exists_{1/2} y R(x,y) \\ x \notin L & \implies \exists_{1/2} y \neg R(x,y) \end{cases}$$

Τυχαιότητα (Randomness) VII

- Λόγω της έλλειψης φράγματος για την πιθανότητα λάθους, δεν μπορούμε να χρησιμοποιήσουμε την τεχνική της επανάληψης για να βελτιώσουμε την πιθανότητα σωστού αποτελέσματος από έναν PP αλγόριθμο. Μία άλλη ένδειξη για το ανέφικτο της κλάσης PP σε σχέση με τις BPP, RP, ZPP, προκύπτει από το παρακάτω αποτέλεσμα:

Πρόταση

$\mathrm{NP} \subseteq \mathrm{PP}$.

- Πρέπει επίσης να σημειώσουμε ότι δεν λάβαμε καθόλου υπ' όψιν μας, ως υπολογιστικό πόρο, των αριθμό των τυχαίων bits που χρησιμοποιεί ένας πιθανοκρατικός αλγόριθμος. Στην πράξη, κάθε «τυχαίο» bit που χρειαζόμαστε δεν είναι χωρίς τίμημα, αφού το λαμβάνουμε από κάποια γεννήτρια *ψευδοτυχαίων* bits.
- Τέλος, αναφέρουμε και την κλάση RL (**Randomized Logspace**) που περιέχει τα προβλήματα που έχουν one-sided error αλγόριθμο που χρησιμοποιεί λογαριθμικό χώρο και πολυωνυμικό ως προς το μήκος της εισόδου αριθμό τυχαίων bits.