**Εργαστήριο Αλγορίθμων 1**

Η εταιρία MS-detect με κύρια δραστηριότητα σε υπηρεσίες «εντοπισμού» και «κατάταξης» κακόβουλων λογισμικών (malware), θέλει να αναπτύξει ένα σύστημα το οποίο, δοθέντος ενός λογισμικού Τ, να μπορεί να αποφανθεί εάν πρόκειται για κακόβουλο λογισμικό ή όχι (malware or benign).

Η εταιρία έχει στη διάθεσή της 48 οικογένειες από κακόβουλα λογισμικά (malware). Θεωρείστε για ευκολία, ότι κάθε οικογένεια αποτελείται από ένα μέλος. Κάθε μέλος μιας οικογένειας έχει ορισμένα χαρακτηριστικά από το σύνολο *ch* = {1, 2, …, 10}, όπου:

1: Environment 6: Time

2: Memory 7: Object

3: Registry 8: Job

4: Security 9: Device

5: Process  10: Other

1) Σχεδιάστε και αναλύστε έναν αλγόριθμο ο οποίος θα παίρνει ως είσοδο ένα κακόβουλο λογισμικό *Τ*1 (malware), με *k* χαρακτηριστικά από το σύνολο *ch*, 1 ≤ *k* ≤ 10, και θα κατατάσσει το *Τ*1 σε μία από τις 48 οικογένειες σύμφωνα με τα χαρακτηριστικά του και τα χαρακτηριστικά των οικογενειών, χρησιμοποιώντας μια μετρική ομοιότητας.

2) Σχεδιάστε και αναλύστε έναν αλγόριθμο ο οποίος θα παίρνει ως είσοδο ένα αγνώστου τύπου λογισμικό *Τ*2 (malware ή benign), με *k* χαρακτηριστικά από το σύνολο *ch*, 1 ≤ *k* ≤ 10, και θα αναγνωρίζει τον τύπο του, δηλ. εάν το *Τ*2 είναι κακόβουλο ή όχι. Το λογισμικό *Τ*2 είναι κακόβουλο εάν έχει περισσότερο από 80% ομοιότητα με ένα τουλάχιστον μέλος από τις 48 οικογένειες.

Στο αρχείο lab1std.java που σας δίδεται να συμπληρώσετε το κατάλληλο κώδικα, ο οποίος θα κατατάσσει το λογισμικό *Τ*1 στην κατάλληλη οικογένεια, και θα αναγνωρίζει εάν το λογισμικό *Τ*2 είναι κακόβουλο ή όχι. Το πρόγραμμά σας θα πρέπει να επιστρέφει κατάλληλα μηνύματα, όπως:

* Το κακόβουλο λογισμικό *Τ*1 ανήκει στην οικογένεια 42, ή
* Το λογισμικό *Τ*2 είναι κακόβουλο και ανήκει στην οικογένεια 3, ή
* Το λογισμικό *Τ*2 δεν είναι κακόβουλο.