

### 3<sup>ο</sup> Σύνολο Ασκήσεων

**Καταληκτική Ημερομηνία Παράδοσης:** Τετάρτη 29 Μαΐου 2013, στην αρχή του μαθήματος  
**Ύλη:** Κεφάλαια 19 (εκτός του 19.6), 20 και 21 από το βιβλίο  
*Μπορεί να γίνει προφορική εξέταση*

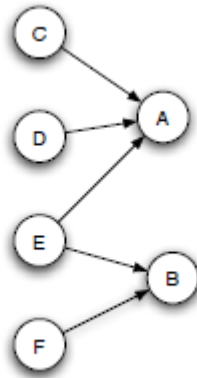
#### Άσκηση 1

Το παρακάτω ερώτημα αναφέρεται στον τρόπο υπολογισμού του μεγέθους του (indexable) web εδώ <http://www.worldwidewebsite.com/>

- (α) Ποια από τις προσεγγίσεις που αναφέρονται στο κεφάλαιο 19.5 χρησιμοποιεί για δειγματοληψία (sampling) και πως κάνει τον έλεγχο (testing);  
(β) Πως αντιμετωπίζει κάποια από τα προβλήματα/θέματα αυτών των μεθόδων

#### Άσκηση 2

Θεωρείστε το παρακάτω γράφο



(α) Δώστε τις τιμές hub και authority μετά από 2 επαναλήψεις του υπολογισμού τους. Δώστε τις τιμές πριν και μετά την κανονικοποίηση, δηλαδή τη διαίρεση τους με το άθροισμα των hub και authority τιμών αντίστοιχα.

(β) Θέλουμε να δημιουργήσουμε μια νέα σελίδα X και να την προσθέσουμε στο γράφο ώστε αυτή να έχει το μεγαλύτερο δυνατό (κανονικοποιημένο) authority score. Μαζί με τη X σκεφτόμαστε να προσθέσουμε και μια δεύτερη σελίδα Y που να δείχνει στη X για να της αυξήσει το score. Εξετάζουμε δύο επιλογές:

Επιλογή 1: Προσθέτουμε δύο νέους κόμβους X και Y στο Σχήμα 1 και δημιουργούμε μόνο ένα σύνδεσμο από το Y προς το X

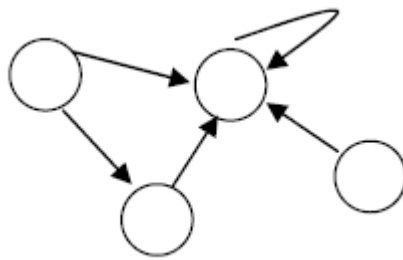
Επιλογή 2: Προσθέτουμε δύο νέους κόμβους X και Y στο Σχήμα 1 και δημιουργούμε συνδέσμους από το Y προς το καθένα από τα A, B και X

Για κάθε μια από τις 2 επιλογές, υπολογίστε μετά από 2 επαναλήψεις το κανονικοποιημένο authority score των A, B και X. Για ποια από τις 2 επιλογές, το X παίρνει καλύτερο score; Δώστε μια σύντομη διαισθητική εξήγηση γιατί.

(γ) Τώρα υποθέστε ότι αντί για δύο νέους κόμβους, δημιουργείτε τρεις κόμβους X, Y και Z και πάλι θέλετε να τοποθετήσετε συνδέσμους από αυτούς ώστε ο X να παίρνει το υψηλότερο δυνατό score. Περιγράψτε μια στρατηγική που το πετυχαίνει και δείξτε πως μετά από 2 επαναλήψεις, ο κόμβος X έχει το δεύτερο καλύτερο κανονικοποιημένο authority score από όλους τους κόμβους.

### Άσκηση 3

Θεωρείστε το παρακάτω γράφο.



Υπολογίστε το PageRank των κόμβων του με και χωρίς τηλεμεταφορά με πιθανότητα 0.1.