Homework 1

(Information retrieval, Multimedia)

Due: 4:00pm, Feb. 20

1. Write every step in detail.

   a) Build a search engine for the following documents:

   Doc1: apple tree, organic, producer, fruit,

   Doc2: apple, orchard, cooking, producer, fruit, US

   Doc3: apple iphone, US, app,

   b) show the retrieved results when doing the following queries:

   apple, apple & fruit, apple –fruit, apple & US

   where,

   "&" means two words are both used in the documents

   "-fruit" means "fruit" should not be in the retrieved documents

2. Design an image retrieval system based on:

   A collection of images, each image has several tags.

   This retrieval system should be able to handle:

   1) user can get desired images using key words

   2) user can get desired images using a sample (query) image

   3) user can get what are the objects on the image

   Write every step in detail to show your design and show why you can handle these requests?

   (hint: you should mention how to build text based retrieval system, how to build image based retrieval system, how to build annotation system, using concepts like: sample, feature, label, training set, test set, model, clustering, classification, supervised, unsupervised)