Computer Engineering Department
Bilkent University

CS533: **Information Retrieval Systems**
Assignment No. 3. Version No. 2 of November 7, 7:12 pm.
October 27, 2017
Due date: November 9, 2017; Thursday, class time (hardcopy is required).

Note: Handwritten answers are not acceptable. ~~This is a progressive assignment: more questions will come.~~

1. **Motivation**: Consider finding the number of clusters problem in C3M that involves the summation of the diagonal elements of the C matrix. As we find the summation we add small numbers to a progressively becoming large number and due to IEEE 754 floating point representation we expect to obtain an incorrect results since as we add a small number to a large number the increase provided by the small number can be lost.
   Experiment - Write a Program: Generate a floating point number array of size 2^{20}. Initialize the array elements by random numbers between 0 and 0.0000001 (10^{-7}). This is Case 1.

   a. Find the summation of the array elements by doing the summation 1 by 1.
   b. Find the same summation by adding the elements 2 by 2. This means adding the intermediate results in the second and later stages. For example if we have 8 integer numbers: 1, 2, 3, 4, 5, 6, 7, 8; first stage generates 3, 7, 11, 15, second stage generates 10, 26, and the third and the final stage generates 36
   c. Find the summation 4 by 4 (2^2 by 2^2), n= 2.
   d. Find the summation 2^3 by 2^3, n= 3.
   e. Do the same for n= 4, 5, ..., 19. Note that n= 20 is the same as doing the summation 1 by 1.
   f. Case 2: Repeat the above experiment for random numbers between 0 and 0.00000001 (10^{-8})
   g. For Case 1 and Case 2 obtain the percentage of error. Make sure that you define percentage of error. Give a plot for this. Two lines in the same figure. The x axis is for n varying from 1 to 20. The y axis is for the percentage error.
   h. Write a short report for this experiment and state your conclusion based on your observations. If you are unable to conclude anything state why.
   i. If you like design another experiment for this problem and give your results for you design.