Due: By 12:00 noon on Thursday, November 16. **Worth**: 7%
You must complete and sign an assignment cover page, and attach it (with a staple) to the front of your assignment. Assignments should be handed into the drop box in BA 2220.

1. [8 marks] Consider the algorithm in Question 4 from the previous assignment. We assumed that the while loop eventually terminated, but didn't argue why this assumption would be true.

   (a) Suppose there are n beans at the beginning of an iteration of the while loop. How many beans are at the end of this iteration? Express your answer symbolically.

   (b) Prove your answer to (a).

   (c) Using your answer to (a), give a justification that the while loop will terminate.

2. [15 marks] Using our structured form, prove or disprove each of the following:

   (a) $n^4 - 5n^3 - 5n^2 - 5n \in O(n^4)$

   (b) $n \log_2(n) \in O(n)$

   (c) $\frac{4n^3}{6n^5+3n^3} \in \Theta(\frac{1}{n^2})$

3. [10 marks]

   Let $\mathcal{F}$ be the set of functions from $\mathbb{N}$ to $\mathbb{R}^{\geq 0}$. Prove the following theorems about asymptotic notation:

   (a) For $f, g \in \mathcal{F}$, if $g \in \Omega(f)$ then $g^2 \in \Omega(f^2)$.

   (b) For $k \in \mathbb{N}$, $k > 1 \Rightarrow \forall d \in \mathbb{R}^+, d \log_k n \in \Theta(\log_2 n)$.

4. [7 marks]

   Data compression is commonly used to try to reduce the amount of memory required to run programs.

   Let's say that we are trying to represent DNA sequences in an array of 16 bit integers. DNA sequences are comprised of four nucleotides, A = adenine, C = cytosine, G = guanine, T = thymine. A simple method to represent the sequence is to represent each nucleotide as an integer, so A = 0, C = 1, G = 2, T = 3, and then store these integers in the integer array.

   So if we are trying to represent the DNA sequence 'CGCA', the integer array will have four elements: 1 2 1 0. The bit representation of this integer array is:
   0000000000000001 0000000000000010 0000000000000001 0000000000000000.

   (a) Suppose the DNA sequence has n nucleotides. How long does our integer array have to be to store this sequence?

   (b) Devise and describe a storage scheme where you can represent an n nucleotide chain in an $\frac{n}{8}$ element array of 16 bit integers. Show how to look up the i^{th} element of the nucleotide chain in constant time.

   (c) Now let's generalize our compression scheme. Suppose our domain of things to store in our array consists of k different elements, and our integers are m bits long. How many items could we represent in a single m bit integer?