Lab Sheet 2

This lab is a warmup: it will not be collected. It involves some simple exercises using Python's regular expression machinery.

1. Write a function `moolah` that takes a string as its argument and that returns a list of every Euro amount (as a string) that is mentioned in the text. We assume that Euro amounts appear as follows: (i) the acronym EUR, (ii) optionally followed by a space, (iii) followed by one or more digits, (iv) optionally followed by a decimal point and one or more digits. The following are all valid examples.

   ```
   EUR100    EUR 250   EUR 12.50 EUR1000000
   ```

   Note our definition disallows negative amounts and the use of commas or spaces to separate clusters of digits.

2. Write a function `bleep` that takes a string as its argument and that returns a modified copy of the text in which all four-letter words are replaced by `****`.

3. Write a function `to_english` that takes a string as its argument and that returns a modified copy of the text in which all numbers appearing in the original are spelled out in English. For example,

   ```
   My favourite number is 17!
   ```

   would appear in the output as

   ```
   My favourite number is 17 (one seven)!
   ```

   You need cope only with non-negative integers.

4. Write a function `harvest_emails` that takes a string as its argument and that a returns a list of the email addresses that occur within that string. The list should be ordered by domain and by local part for those addresses sharing the same domain.

   We will assume that each email address

   $$\underbrace{\text{John.Patrick.Murphy3}}_{\text{local part}}@\underbrace{\text{computer-science.ucc.ie}}_{\text{domain}}$$

   comprises a *local part*, an at symbol (@) and a *domain*. The local part consists of upper or lowercase letters, digits plus the underscore and dot symbols; a dot may not occur as the first or last symbol and no two dots may appear one after the other. The domain consists of a dot separated list of labels, each of which is composed of letters, digits or hyphens. Hyphens may not occur as the first or last symbol in any label.

   Try to avoid searching for a solution online: we need to get away from the idea of relying on looking up solutions online in favour of figuring things out for ourselves.