School of Computing and Information Systems
The University of Melbourne
COMP90049 Knowledge Technologies (Semester 2, 2017)
Workshop exercises: Week 2

1. Observe from the lecture slides that we have definitions (of sorts) for **data** and **information**, but no explicit definition of **knowledge** is given. Reread the lecture slides and revise the definitions of the two terms above, and choose salient points of the lecture to define **knowledge** (perhaps in terms of the other two concepts).

2. What do we mean when we say **knowledge technologies**?

   (a) Revise the definition of **knowledge tasks**, with respect to **concrete tasks**.

   (b) Consider the following, and decide into which category you believe they fall, referring to the definition you have decided upon above.

      i. Multiplying two floating-point numbers in base 16
      ii. Playing a competitive game of naughts-and-crosses
      iii. Playing a competitive game of go
      iv. Playing a competitive game of tennis
      v. Calculating the trajectory of a thrown book
      vi. Selecting appropriate counter-measures after someone has thrown a book at you
      vii. Selecting a book that a given person will enjoy reading
      viii. Translating a program written in C into Java
      ix. Translating a document written in Japanese into English

3. What are **structured data** and **unstructured data**? Give an example of each, and indicate how you would handle each in a computational setting.
   What about **semi-structured data**? Are any of the examples you gave above actually instances of semi-structured data? In what ways is it easier or more challenging to handle semi-structured data?

4. Describe a process through which we might be able to answer the question "Where shall we go for dinner tonight?" using Google (`http://www.google.com`) as a resource.
   (We'll touch on some of these elements as the semester goes on.)

5. Revise the following **regular expression** operators:

   ```
   ( ) [ ] { } . * + ? ^ $ | \
   ```

   For each of the following, give a couple of examples of strings which the regular expression would match. Describe (colloquially, in a manner that a non-technical person would understand) the set of strings that the pattern is designed to match.

   (a) `/[a-zA-Z]+/`
   (b) `/^[A-Za-z][a-z]*$/`
   (c) `/p[aeiou]{,2}t/`
   (d) `/\s(\w+)\s\1/`

6. Write a regular expression to solve the following problems:

   (a) Match a price
   (b) Match an Australian telephone number
   (c) Remove HTML comments from a document
   (d) Validate an email address