**Upload your project report and related documents in MyCourses by 11:59PM, 12/03/2021**

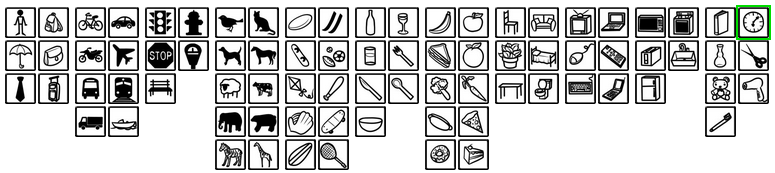
This project is a teamwork project. You can complete this project with your team members. Each team cannot have more than 3 team members. But if you want to complete this project by yourself, it is fine. Your team can use any programming language. For each team or individual submission, please submit the work by only one team member. Please do not submit duplicate copies of the team work.

**Project 3. Construct Knowledge Graph Using COCO Database**

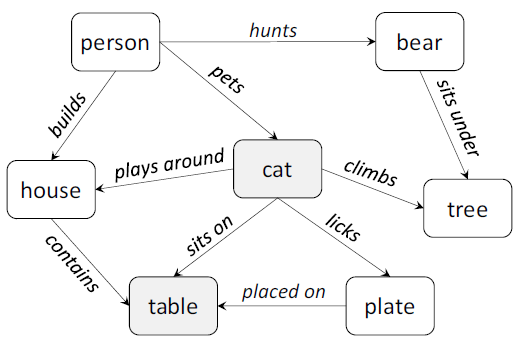
Please read the paper “Object Detection Meets Knowledge Graphs”. This paper is uploaded with this project instruction document on MyCourses. You can find it from the MyCourses Project 3.

Based on your and your team members’ understanding, implement the partial content of the paper “Object Detection Meets Knowledge Graphs” using COCO database. The following steps can be used for achieving the project goals.

**Step1:** Go to <https://cocodataset.org/#explore>, and download 100 images for building your image database. For example, if input the “table” for exploration, many images are displayed. Select 100 images and download them to your own computer. You can choose any one keyword such as “cat”, “car”, “bowl”, “apple”, “clock”, and other words included in the COCO database. Most popular concepts are included in the following figure. After you select 100 images, put those images into one folder named “image100”. Extract all concepts from each image and build a concept pool for all 100 images. For example, if the first image has concepts “cat” and “table”, the second image has “bowl” and “banana”, and the third image has “sofa” and “chair”, the concept pool of those 3 images is {“cat”, “table”, “bowl”, “banana”, “sofa”, “chair”}.



**Step2:** After obtaining all concepts of 100 images and building the concept pool, manually build one knowledge graph for one image, which can be selected from those 100 images. For example, the following knowledge graph represent an image, which contains 7 concepts. Each node represents a concept.



**Step3:** Write a computer program to construct the frequency-based knowledge matrix as shown in the subsection 3.2 of the paper “Object Detection Meets Knowledge Graphs”. This frequency-based knowledge matrix has the size of all concepts extracted in the **Step1**. For example, if 196 concepts are extracted from 100 images in the **Step1**, this frequency-based knowledge matrix’s size will be 196 x 196. Using the Equation (1) in the paper “Object Detection Meets Knowledge Graphs”, each element value of this matrix can be calculated.

**Step4:** Based on semantic consistency generated from the frequency-based knowledge matrix in the **Step3**, write a program to retrieve the 10 concepts’ relations with three highest frequencies. For example, if “cat” and “table” has the co-occurrence value 1.3 in the frequency-based knowledge matrix, “cat” and “sofa” has the co-occurrence value 1.1, and “cat” and “chair” has the co-occurrence value 0.9, the retrieval results are presented in the second row of the following table. Please list 10 concepts (you can choose any 10 concepts extracted from 100 images in the **Step1**) and their 3 strongest relations in the following table. Copy this table into your project report in the **Step5**.

|  |  |  |  |
| --- | --- | --- | --- |
| Concepts (1-10) | Relation 1 | Relation 2 | Relation 3 |
| **cat** | table | sofa | chair |
| **apple** | plate | bowl | knife |
| … | … | … | … |
| … | … | … | … |
| … | … | … | … |

**Step5:** Write your project report by answering the following questions.

(5.1) How many concepts in 100 images you collected? List all concept names in the project report.

(5.2) What programming language does your team use in this project?

(5.3) List the table of the Step4 in the project report.

**Upload the following documents (compressed all files in a zip file) on MyCourses by 11:59PM, 12/03/2021 for grading evaluation. This project is a teamwork project (no more than 3 team members), but you can also complete this project by yourself.**

1. **Upload your 100 COCO images folder named “image100” (in a zip file) acquired by your team in the Step1 to MyCourses.**
2. **Upload one knowledge graph you manually generated in the Step2.**
3. **Upload the computer program for implementing the Step3 and Step4.**
4. **Upload your project report in the Step5.**