Assignment 5: Image Description Generation

Due: December 2

Check the following multi-media Information Extraction output example about protest events (Please use Google Chrome or Safari):
http://nlp.cs.rpi.edu/multimedia/Demonstrate.html

For each cluster, you will see the following rows:
1. images sharing similar visual patterns
2. part of images indicating the visual patterns
3. Ground-truth text captions
4. Text IE results
5. New event types: Verb concept clusters discovered by text IE
6. New noun concept clusters discovered by text IE

What you need to do:

a. Please implement a basic natural language generation system (e.g., use ngrams or part-of-speech tagging) to compose 4-6 above to form a natural language sentence; You can write a code to do this, or write templates manually. (7pt)

b. Test the system on 10 clusters in the data, compare the results with ground-truth captions, and analyze errors, propose new ideas for improvement; (3pt)

c. Implement new ideas beyond baseline approach. (5 extra pt)