WebSem – Spring 2020
Introduction to the Semantic Web

Part 1: Ontology Modeling and Logic

1. The constructs the constitute ALC are intersection, union, complement, universal restriction and existential restriction.
2. TBox and ABox
    a. The TBox will contain the collection of inclusion and equivalence between concepts, i.e. $\mathsf{marry}(x,y) \equiv \mathsf{loves}(x,y) \cap \mathsf{loves}(y,x)$. The ABox will contain the assertions applied on the individuals, i.e. friend(John,Susan) AND friend(John,Andrea) AND loves(Susan,Andrea) AND loves(Andrea,Bill) AND Female(Susan) AND NOT Female(Bill) AND age(John,22).
    b. No, John does not have a female friend loving a male person since we do not know the gender of Andrea. To solve this problem, we can add Female(Andrea).
3. An object property relates entities between them, e.g. SubClass, EquivalentClass, etc. and a data type property relates instances to literal data such as strings, numbers, etc. An annotation property is used to give a name to instances, classes or properties.
4. A property is defined as a mapping between two sets of instances (an oriented link in the knowledge graph). So, to define the domain and the range of a property is to define respectively the starting and the resulting class of each property. Age could have Person as domain and Integer as range.
5. Universal: having only one relationship along this property, and existential: participating in many relationships using this property.
6. Assertions
    a. $dbo{:}Writer \sqsubseteq dbo{:}Person$
    b. $m{:}Book \sqsubseteq= 1.has.m{:}publicationYear$
    c. $m{:}Novel \sqsubseteq m{:}Book \sqsubseteq m{:}Work$
    d. $dc{:}title \sqsubseteq= xsd{:}string$

Part 2: Query Knowledge Graph

7.

```
SELECT ?rapid_transit ?rapid_transitLabel ?loc ?date WHERE {
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
  ?rapid_transit wdt:P1192 wd:Q5503.
  ?rapid_transit wdt:P625 ?loc .
  ?rapid_transit wdt:P1619 ?date .
}
LIMIT 100
```

8. None
9.
    a. This table shows philosophers, their names in Wikidata and their birth dates in both Wikidata and DBpedia.
    b. Service is used to specify which knowledge base to use.

Part 3: Information Extraction

10. Tweet
    a. There so many preprocessing tasks to be done: normalization (e.g. remove capitals), language identification, tokenization (split the tweet into atomic parts e.g. words) and part-of-speech tagging (define the grammatical role of each token).
    b. [lakers, lebron, james, became, third-leading, scorer, nba, history, staples, center]
    c. None
    d. Person: LeBron James, Location: Staples Center, Organization: NBA.
11. Use a statistical approach for NER and then link it with KB indexing (using popularity for example).