COURSE: NATURAL LANGUAGE PROCESSING

ASSIGNMENT

(Course group:CC01, Term: 2022-2023)

Topic: Build a simple question-answering system for domestic flights

1. Given database of flights

(FLIGHT VN1), (FLIGHT VN2), (FLIGHT VN3), (FLIGHT VN4), (FLIGHT VN5)

(FLIGHT VJ1), (FLIGHT VJ2), (FLIGHT VJ3), (FLIGHT VJ4), (FLIGHT VJ5)

(ATIME VN1 HUE 11:00HR) (DTIME VN1 HCMC 10:00HR)

(ATIME VJ1 HUE 13:30HR) (DTIME VJ1 HN 12:30HR)

(ATIME VN2 HCM 16:30HR) (DTIME VN2 ĐN 15:30HR)

(ATIME VJ2 HN 11:00HR) (DTIME VJ2 ĐN 9:30HR)

(ATIME VN3 HN 6:30HR) (DTIME VN3 HCM 4:30HR)

(ATIME VJ3 HP 11:45HR) (DTIME VJ3 HCMC 9:45HR)

(ATIME VN4 ĐN 11:30HR) (DTIME VN4 HN 9:30HR)

((ATIME VJ4 ĐN 9:30HR) (DTIME VJ4 HCMC 8:30HR)

(ATIME VN5 KH 17:45HR) (DTIME VN5 HCMC 17:00HR)

(ATIME VJ5 KH 10:45HR) (DTIME VJ5 HN 9:00)

(RUN-TIME VN1 HCMC HUE 1:00 HR) (RUN-TIME VJ3 HCM HP 2:00 HR)

(RUN-TIME VJ1 HN HUE 1:00 HR) (RUN-TIME VN4 HN ĐN 2:00 HR)

(RUN-TIME VN2 ĐN HCM 1:00 HR) (RUN-TIME VJ4 HCM ĐN 1:00 HR)

(RUN-TIME VJ2 ĐN HN 1:30 HR) (RUN-TIME VN5 HCM KH 0:45 HR)

(RUN-TIME VN3 HCM HP 2:00 HR) (RUN-TIME VJ5 HN KH 0:45 HR)

Note: explanation for symbols in the database:

1.1. (DTIME VN1 HCMC 10:00HR) contains following elements:

- DTIME: departure time of flight
- VN1: Việt Nam airline flight code is VN1;
- HCMC: Hồ Chí Minh city
- 10:00HR: departure hour is 10:00HR

1.2. (ATIME VJ1 HUE 13:00HR) contains following elements:

- ATIME: arrival time of flight ,
- VJ1: VietJetAir flight code is VJ1

- HUE: Huế city
- 13:00HR: arrival hour is 13:00HR

1.3. (RUN-TIME VN2 ĐN HCM 1:00HR) contains following element:

- RUN-TIME: flight time,
- VN2: flight code is VN2,
- ĐN: Đà Nẵng city is the place, where flight departs.
- HCM: Hồ Chí Minh city is the place, where flight arrives.
- 1:00HR: flight time from departure place to arrival place

2. Requirement

2.1 Given Vietnamese queries

1) Máy bay nào đến thành phố Huế lúc 13:30HR ?.
2) Máy bay nào bay từ Đà Nẵng đến TP. Hồ Chí Minh mất 1 giờ ?.
3) Hãy cho biết mã hiệu các máy bay hạ cánh ở Huế ?.
4) Máy bay nào xuất phát từ Tp.Hồ Chí Minh, lúc mấy giờ ?.
5) May bay nào bay từ TP.Hồ Chí Minh đến Hà Nội ?.
6) Máy bay VN4 có xuất phát từ Đà Nẵng không ?.
7) Thời gian máy bay VJ5 bay từ TP. Hà Nội đến Khánh Hòa mất mấy giờ ?.
8) Có máy bay nào xuất phát từ Hải Phòng không ?.
9) Máy bay của hãng hàng không VietJet Air bay đến những thành phố nào ?
10) Có máy bay nào bay từ Hải Phòng đến Khánh Hòa không?

2.2 Write a program to implement the above Vietnamese queries

The program implements following functions:

a) Build a dependency parser.
b) The dependency parser can create relations for elements of each query.
c) From results of b), construct *grammatical relations* for flights between mentioned cities
d) Create logical forms from *grammatical relations* of c)
e) Create procedural semantics from logical forms of d) with the given database.
f) Retrieve the database to search appropriate answering information for mentioned queries.

3. Request for implementation

3.1 ***Programming language is*** Python. Entry point is **main.py,** which is in the origin directory.

Student's assignment will be marked on Python 3.7.

3.2 For ***output***

For each query a, b, c… student should create the output file with the name in order such as: output_a.txt, output_b.txt, output_c.txt, …

3.3 ***Submission of the assignment***

- Student must compress all of files/directories (source code, data,…) into compressed file with form **MSSV.zip**, in which MSSV is student code. Student does not compress the file into other forms. Note that, when student decompress the file, then directories and 2 files will appear, as follows.
  - - Each student has 3 directories:
  - Input: input files are file.txt of the topic (queries, database,…)
  - Output: implementation result of each query in the part 2.1 is output of previous request (example: b), then it is input for next request (example: c)
  - Models: directory that contains classes or sub-modules, which implement the topic.
- In the directory of each student, one file **README.md** contains personal information and notes for modules or classes in the directory Modes
- In addition, there is one main.py according to the request of a).

3.4 Request: the program when execution is without parameter as follows: **python main.py**
