

The project aims at extracting the location of the observed objects from astronomy circulars in an automated way. This would be used to create a page to query all the circulars which referred to a particular location.

Professor Varun Bhalerao has suggested the following steps to proceed with the problem.

- 1) Identify the line which gives information about coordinates because it can be known from the context
- 2) Identify the numbers which are coordinates
- 3) Convert the coordinates into the standard form which can be done via a switch case (without NLP)

Following are the examples discussed with Professor Bhalerao which illustrate a few examples:

1)<http://www.astronomerstelegam.org/?read=13354>

Inference:

The first paragraph gives the coordinates ie “ R.A.: 60.43819 deg, Dec.: 21.17461 deg ”

2)<http://www.astronomerstelegam.org/?read=13361>

Inference:

There is no explicit mention of coordinates however the coordinates can be extracted from the object name ie “J060000.76-310027.83” in the first line of the first paragraph. From this, it can be inferred that the object has Ra = 06h 00min 00.76sec and Dec = - (31deg 00 min 27.83sec). However in the first example also we had such object names but since explicit coordinate names are more accurate than those inferred from object names, we should report that.

3)<http://www.astronomerstelegam.org/?read=13351>

Inference:

This has no information about coordinates explicitly or implicitly.

4)<http://www.astronomerstelegam.org/?read=13330>

Inference:

This has many coordinates given in the tabular format and all need to be extracted.

5)<http://www.astronomerstelegam.org/?read=13347>

Inference:

This also has no coordinates but illustrates that if we use regex there will be false matches like “2019-12-06 13:10:19 to 2019-12-07 09:28:36 (UTC)” from the first sentence of the first para.