



## Experiment Two

### *Solr*

518030910279 qubowen

April 20, 2019

### First Configure schema

The first step is creating a solr core of Paper. I create a dictionary named 'Paper' and copy the folder 'conf' to the dictionary. Then I create core on solr admin which name and instanceDir are Paper.

The next step is add the required seven schema by 'add field' option. The fields 'PaperID', 'Title', 'ConferenceID', 'ConferenceName' and 'Year' are added whose field type is 'text\_en' in order to fuzzy matching, 'stored' is true to display in the searching results, 'indexed' is true to build index of these items, 'required' is true for every paper has these items. However, for 'Authors\_Name' and 'Authors\_ID', 'multiValued' is true for a paper may have many authors, 'required' is false, and the field type of 'Authors\_Name' is set to 'string' for it needs precision matching. Here we take field 'Title' for an example. Figure 1 and figure 2 is the operation.

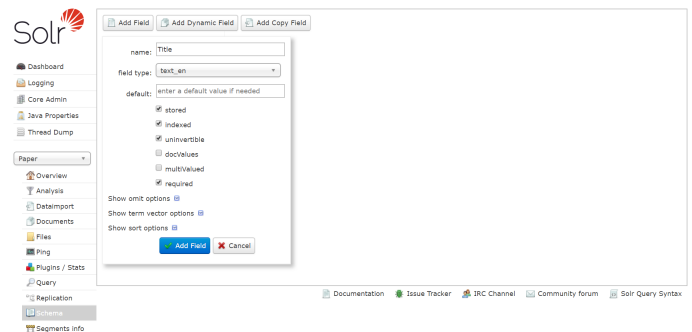


Figure 1: configure of the field 'Title'.

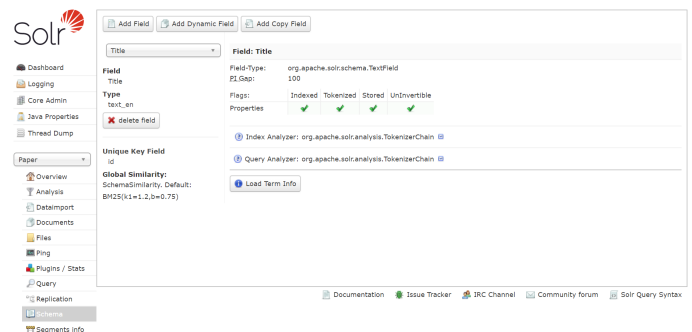


Figure 2: the field 'Title'.

## Second Section-Solve problems

Add data.

I use python with the library 'pysolr' and 'pymysql' to add data from mysql to solr. In figure 3 on the left is my code. First in line 1 and line 2 import library 'pysolr' and 'pymysql'. Then in line 3 connect python to solr. In line 4 clear the data in the solr core 'Paper'. From line 5 to line 10 connect mysql to python. In line 11, create the cursor of the database. Then in line 12 to line 13 download all the papers data to the variable 'results' in type 'list'. In line 14 build a list papers to store the data which will be commit to solr. Then in line 15 to 27, for each paper in 'results', create dictionary paper to store 'PaperID', 'Title', 'Year' and 'ConferenceID' of the paper in results, then select 'ConferenceName', 'Authors\_ID' and 'Authors\_Name' from the database and add it to the paper dictionary. Especially for 'Authors\_ID' and 'Authors\_Name' may have several values, so I build a list to store it. In line 28 add the papers data to the solr and in line 29 commit data to solr.

### Problem 1.

Problem 1 is to search title in solr. First in line 1 import library 'pysolr' and in line 3 connect solr core 'Paper'. Then in line 5 create a variable 'results' to store the results of searching paper from papers whose title contains word 'commonsense' and in line 6 print the results. We can see in the console return all the papers data of all seven attributes whose title contains 'commonsense'.

### Problem 2.

Problem 2 is to search Authors\_Name in solr. First in line 1 import library 'pysolr' and in line 3 connect solr core 'Paper'. Then in line 5 create a variable 'results' to store the results of searching paper from papers whose Author's name is 'kathleen dahlgren' and in line 6 print the results. We can see in the console return all the papers data of all seven attributes whose author's name is 'kathleen dahlgren'..

### Problem 3.

Problem 3 is to search title in solr. First in line 1 import library 'pysolr' and in line 3 connect solr core 'Paper'. Then in line 5 create a variable 'results' to store the results of searching paper from papers whose conference name is 'IJCAI' and in line 6 print the results. We can see in the console return all the papers data of all seven attributes whose conference name is 'IJCAI'.

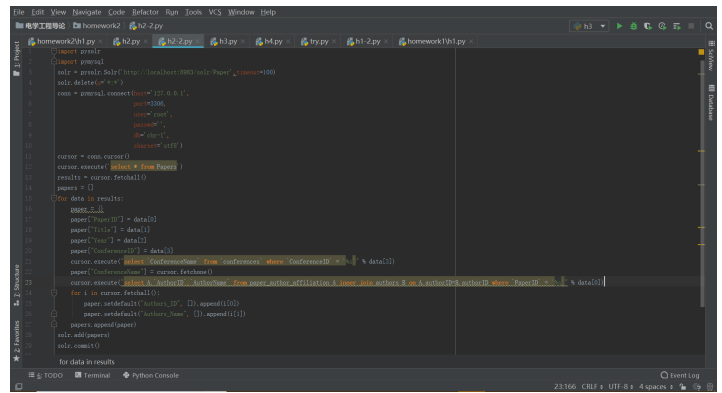


Figure 3: add data from mysql to solr.

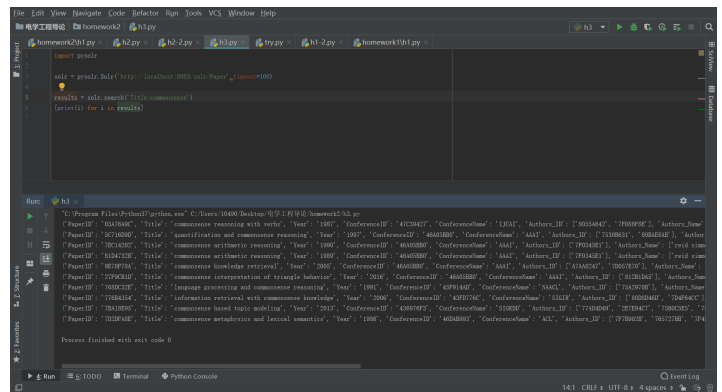


Figure 4: Problem 1:searching Title.

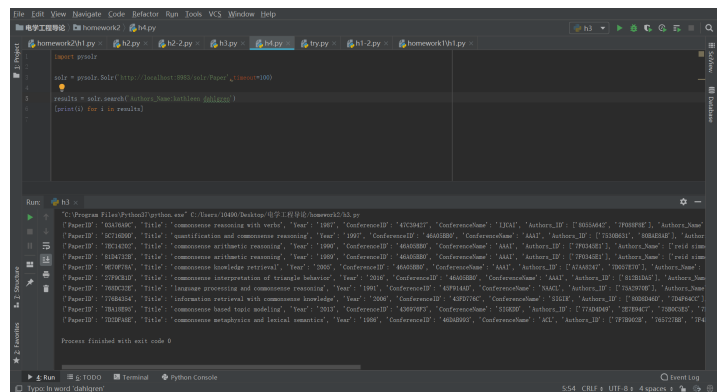


Figure 5: Problem 2:searching Authors Name.

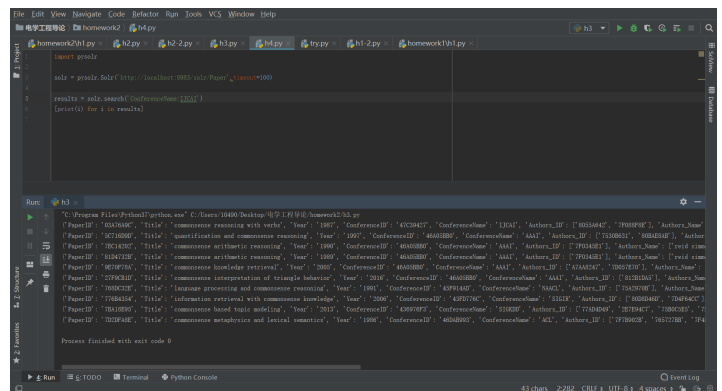


Figure 6: Problem 3:searching ConferenceName.