CCG SPF Documentation

Each table has a unique Primary key. Make the primary key of each table a number.

Names like name of state will come only in one table, in rest of the tables their ids (primary keys will come, helps for cross table referencing)

Predicates for each column

<pkey,t> :- takes primary key and returns truth value : Checks whether we are in correct table or not

<pkey,col\_value> :- takes primary key of a table and returns value of column corresponding to that primary key.

<pkey,<col\_value,t>> :- takes primary key of a table and the column value and returns true if that column value corresponds to that primary key

<pkey,<t,col\_value>> :- takes primary key of a table and the truth value (based on some conditions coming from some part of sentence) and will return col\_value corresponding to primary key if truth value returned is true otherwise returns nothing.

Other predicates

Argmax,argmin, sum, average etc

Pkey\_returner:<col\_value,pkey> :- takes a column value and returns the primary key corresponding to that value. (has type specification hence knows what column to look at)

Set\_returner:<pkey,<col\_value,t>> :- mainly used for conditioning, like in argmax, as argmax takes one function which takes an entity and returns truth value.

First predicate for each column has similar use.

Seed file :-

For each column :- for pred 1 we have type N associated with it.

For type 2 we have NP/NP takes a noun phrase ahead, and NP/PP :- prepositions generally returns primary keys and that will be taken by pred2.

We have N as well as geoquery also has that.

For type 3 we have (S\NP)/NP same was in geoquery dataset

For type 4 we have NP/PP :- here preposition set\_returner function will be used, and it will return truth value which will be taken by the type 4 predicate.

We also have NP/(S\NP) :- here verbs will take something ahead of it and return truth value, which will be taken by the type 4 predicate.

For verbs we have:-

(S\NP)/NP :- type 3 predicate of column used

NP/NP :- type 2 predicate of column used

Synonyms have same lexical entries associated with them as columns.

Adjectives have 3 level (Specific)

Level 1 (large) :- use predicate 2 of columns (NP/NP)

Level 2 (larger) :- PP/NP: use greater and predicate 2 corresponding to that column

Show me the states larger than Gujarat.

Level 3 (largest) :- NP/N: argmax here

(General) largest state by area (Need to have a look at other ways of writing this kind of sentence like largest state in terms of area)

Larger :- PP/NP/NP : Show me states larger than Gujarat by area?

Larger will take first Gujarat after that by area and then state will take that thing as input.

Largest :- NP/N/N: Show me largest state by area

By:- N/N

Preposition

PP/NP :- two functions mainly pkey\_returner and set\_returner.

Can take a function as well which will return some value and then the pkey\_returner will return the primary key corresponding to those values.

Similar thing of taking a function instead of constants can be added to verbs as well.

Code

All predicates and lexical entries and NP list are created in file named predicates. All tose functions are called from generator and generator is called from main. Schema has all the classes required for storing data. Have used Dom parser for reading in XML file.

Schema structure in PPT.

For more clarity can have look at training examples uploaded on github . Code also uploaded on github.

All files generated is uploaded as well.