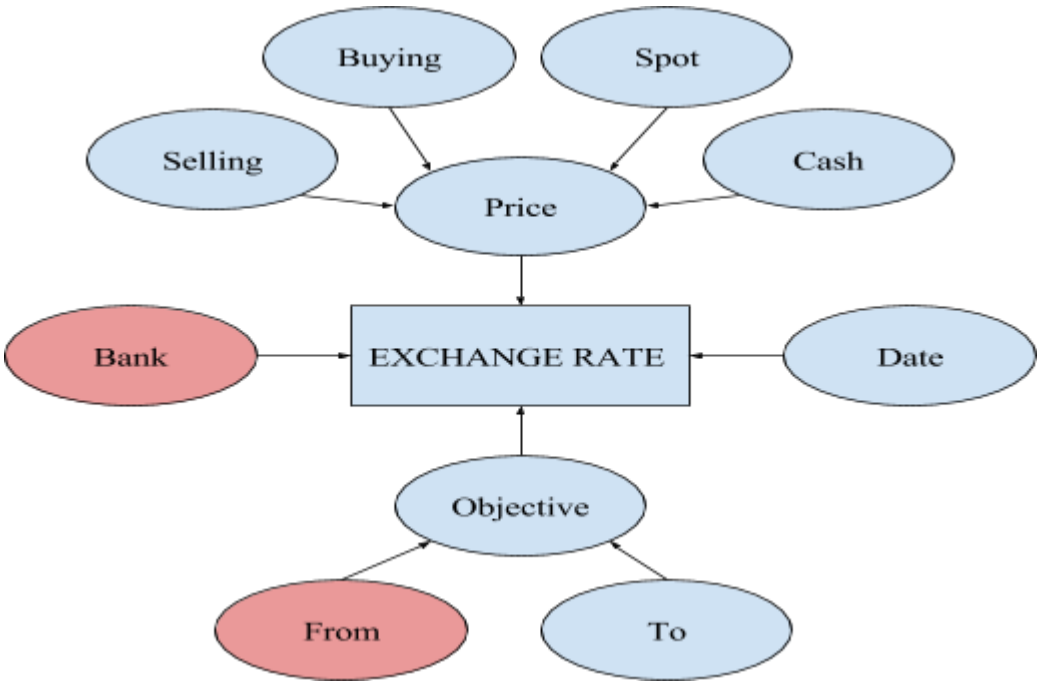
Milestone 1 Ontology and Language Understanding model

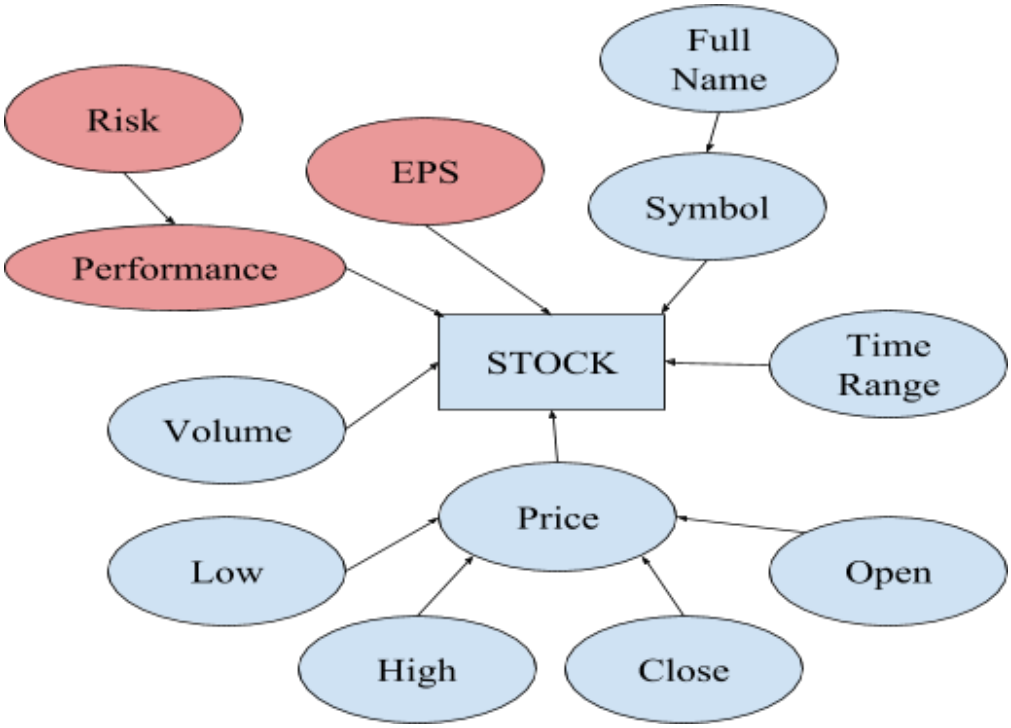
ONTOLOGY

Here shows the graph of our ontology, which can be split into two part: exchange rate and stock. The circles and rectangles with blue background is already in our database, while others with red background will be implemented.

- Exchange Rate



Stock



Backend tables

1. Stock table

symbol	date	open	high	low	close	volume	adj_close
FAX	2016-12-30	4.65	4.70	4.61	4.63	3498900	4.63
IAF	2016-12-30	5.44	5.48	5.40	5.42	103500	5.42

Besides, there is a sub-table for stock table to convert full name to symbol.

symbol	full_name
FAX	Aberdeen Asia-Pacific Income Fund Inc
IAF	Aberdeen Australia Equity Fund Inc

2. Exchange rate table

money_name	cash_sell	spot_sell	cash_buy	stock_buy
USD	30.25	30.55	30.792	30.65
THB	0.7801	0.8686	0.9231	0.9089

Supported functions

1. exchange(‘country1’, ‘country2’)

This is a action for those who want to know the outline of exchange rate between two specific country. The action will return the exchange rate from country 1 to country 2.

2. get_exchange_rate(money_name, buy, type)

The three slots mean goal money, buying/selling and spot/cash respectively. It will return the rate between goal money and NTD. Moreover, the exchange rate here is based on the data from Taiwan Bank.

3. query(name2sym(‘Apple Inc.’), '20170102')

This is the action to get the information of some specific stocks. The first slot is a sub-function that will convert full name to symbol, since the companies are in the form of shorthand in our database. The query action will return the information of some stocks, including open, high, low, close and volume. Now it only supports the data ‘20170102’, we will update the other dates by using the crawler.

Language Understanding Model:

Based on tensorflow technique, I built a bi-directional Recurrent Neural Network model. With the help of pre-trained GloVe word2vec, I am able to transform sentences into word vector. There are still some problems remain unsolved, like how to label OOV and to improve the performance of slot tagging.

Besides exchange action, we have miss tagging or over tagging on the other actions, so the generation of other two still exists some error.

intent	Slot tags			
Query	symbol	date		
Exchange	Country1	Country2		
Get_exchange_rate	Money_name	buy	types	

Training example:

Please give me the exchange rate of FRF to RUB.

O O O O O O O B-country1 O B-country2

Exchange

Testing example:

I like to buy some NZD with cash

O O O O O B-symbol O B-symbol(should be types)

get_exchange_rate

NLG:

implement a class for processing data, convert utters into glove vectors, convert slot tags into one-hot vectors, convert intents into binary vectors.

implement a rule based function for NLG, given action_item, retrieve required information from database and present to user as response.