Training data demonstration

This readme file explains the detail of the data.

The training data of two subtask are *SP-train.npy* and *WP-train.npy*, which shares

same format as following:

```
{'id': 'SP-0_SR',
 'question': 'The six daughters of Mr. and Mrs. Mustard each have one brother. However, the family only consists of nine people; how is that possible?',
 'answer': 'Each daughter shares the same brother.',
 'distrator1': 'Some daughters get married and have their own family.',
 'distrator2': 'Some brothers were not loved by family and moved away.',
 'distrator(unsure)': 'None of above.',
 'label': 1,
 'choice_list': ['Some brothers were not loved by family and moved away.',
 'Each daughter shares the same brother.',
 'Some daughters get married and have their own family.',
 'None of above.'],
 'choice_order': [2, 0, 1, 3]}
```

id represents data type (**SP** for sentence puzzle, **WP** for word puzzle) and data

index.

Question, **answer** represents the brain teaser question and its correct answer.

Distractor represents the incorrectly distractors for our QA problem.

Label represents the true answer index after randomizing the order.

Choice_list contains the choices of QA problem.

Choice_order represents the original question index after randomizing the order.

(**0** means answer, **1** means distractor 1, etc)

Evaluation data demonstration for Codalab

Notice!!! As the current phase on Codalab aim to allow participants to check their submission format. We choose a portion of training data as evaluation data due the limitation in size of whole dataset.

The evaluation data of two subtasks are SP_eval_data_for_practice.npy and WP_eval_data_for_practice.npy, the format are as following:

```
{'question': 'A woman who lives in new york legally married three men, she did not get divorce, get an enrollment, or legally seperate. How is this possible?',
 'choice_list': ['The woman is not a good person.',
  'His husband is not a good husband.',
  'She is a minister.',
  'None of above.']}
```

We only keep question and choice_list as other information like choice_order contains no additional information helping the model performance. Also we remove the id of each data as this can leak answers for some questions. For example, the answer to the original and its semantic reconstruction can be the same.