Embedding Lookup

In [1]:

```
import tensorflow as tf
```

Practical explanation

In the `seq2seq` model we will need to retrieve embeddings for specific words from large tensors that we learn from the model.

This is achieved using the `tf.nn.embedding_lookup` function. In it's simplest form it takes a tensor as the `params` argument and the row `ids` and it returns the values of the tensor at each row represented by the `ids` . For example

In [2]:

```
params = tf.constant([10,20,30,40])
ids = tf.constant([1,0,3,3])
with tf.Session() as sess:
    print(tf.nn.embedding_lookup(params,ids).eval())
```

```
[20 10 40 40]
```

The params argument can be a list of tensors, rather than a single tensor.

In such a case, the indexes, specified in ids, correspond to elements of tensors according to a partition strategy, where the default partition strategy is 'mod'.

In the 'mod' strategy, index 0 corresponds to the first element of the first tensor in the list. Index 1 corresponds to the first element of the second tensor. Index 2 corresponds to the first element of the third tensor, and so on. Simply index i corresponds to the first element of the (i+1)th tensor , for all the indexes 0..(n-1), assuming params is a list of n tensors.

Now, index n cannot correspond to tensor n+1, because the list params contains only n tensors. So index n corresponds to the second element of the first tensor. Similarly, index n+1 corresponds to the second element of the second tensor, etc

Example

See https://stackoverflow.com/questions/34870614/what-does-tf-nn-embedding-lookup-function-do (https://stackoverflow.com/questions/34870614/what-does-tf-nn-embedding-lookup-function-do)

In [3]:

```
params1 = tf.constant([1,2])
params2 = tf.constant([10,20])
ids = tf.constant([2,0,2,1,2,3])
with tf.Session() as sess:
    print(tf.nn.embedding_lookup([params1, params2], ids).eval())
```

```
[ 2  1  2 10  2 20]
```

index 0 corresponds to the first element of the first tensor: 1

index 1 corresponds to the first element of the second tensor: 10

index 2 corresponds to the second element of the first tensor: 2

index 3 corresponds to the second element of the second tensor: 20

See https://stackoverflow.com/questions/34870614/what-does-tf-nn-embedding-lookup-function-do (https://stackoverflow.com/questions/34870614/what-does-tf-nn-embedding-lookup-function-do)

Code

In [4]:

```
import tensorflow as tf
sess = tf.InteractiveSession()
```

Let's say we only have 4 words in our vocabulary: *"the"*, *"fight"*, *"wind"*, and *"like"*.

Maybe each word is associated with numbers.

Word	Number
'the'	17
'fight'	22
'wind'	35
'like'	51

In [5]:

```
embeddings_0d = tf.constant([17,22,35,51])
```

Or maybe, they're associated with one-hot vectors.

Word	Vector
'the '	[1, 0, 0, 0]
'fight'	[0, 1, 0, 0]
'wind'	[0, 0, 1, 0]
'like'	[0, 0, 0, 1]

In [6]:

```
embeddings_4d = tf.constant([[1, 0, 0, 0],
                             [0, 1, 0, 0],
                             [0, 0, 1, 0],
                             [0, 0, 0, 1]])
```

This may sound over the top, but you can have any tensor you want, not just numbers or vectors.

Word	Tensor
'the '	[[1, 0] , [0, 0]]
'fight'	[[0, 1] , [0, 0]]
'wind'	[[0, 0] , [1, 0]]
'like'	[[0, 0] , [0, 1]]

In [7]:

```
embeddings_2x2d = tf.constant([[[1, 0], [0, 0]],
                               [[0, 1], [0, 0]],
                               [[0, 0], [1, 0]],
                               [[0, 0], [0, 1]]])
```

Let's say we want to find the embeddings for the sentence, "fight the wind".

In [8]:

```
ids = tf.constant([1, 0, 2])
```

We can use the `embedding_lookup` function provided by TensorFlow:

In [9]:

```
lookup_0d = sess.run(tf.nn.embedding_lookup(embeddings_0d, ids))
print(lookup_0d)
```

```
[22 17 35]
```

In [10]:

```
lookup_4d = sess.run(tf.nn.embedding_lookup(embeddings_4d, ids))
print(lookup_4d)
```

```
[[0 1 0 0]
 [1 0 0 0]
 [0 0 1 0]]
```

In [11]:

```
lookup_2x2d = sess.run(tf.nn.embedding_lookup(embeddings_2x2d, ids))
print(lookup_2x2d)
```

```
[[[0 1]
  [0 0]]

 [[1 0]
  [0 0]]

 [[0 0]
  [1 0]]]
```

Present Slides Themes Help