In [1]:
```
%pylab inline
import pylab
from sklearn.datasets import fetch_mldata
DATA_PATH = '~/data'
mnist = fetch_mldata('MNIST original', data_home=DATA_PATH)
```

```
Populating the interactive namespace from numpy and matplotlib
```

In [2]:
```
train = mnist.data[:60000]
test = mnist.data[60000:]
```

In [3]:
```
test_sample = test[::50]
```

In [4]:
```
%%time
from sklearn.neighbors import NearestNeighbors
model = NearestNeighbors(algorithm='brute').fit(train)
```

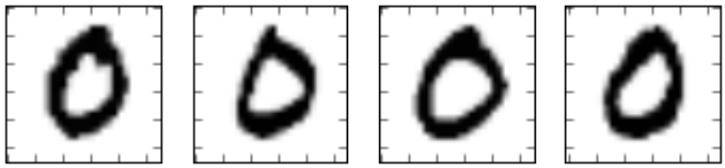
```
CPU times: user 30.8 ms, sys: 14.1 ms, total: 44.9 ms
Wall time: 54.3 ms
```

In [5]:
```
%%time
query_img = test[0]
_, result = model.kneighbors(query_img, n_neighbors=4)
```

```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)

CPU times: user 280 ms, sys: 186 ms, total: 466 ms
Wall time: 443 ms
```

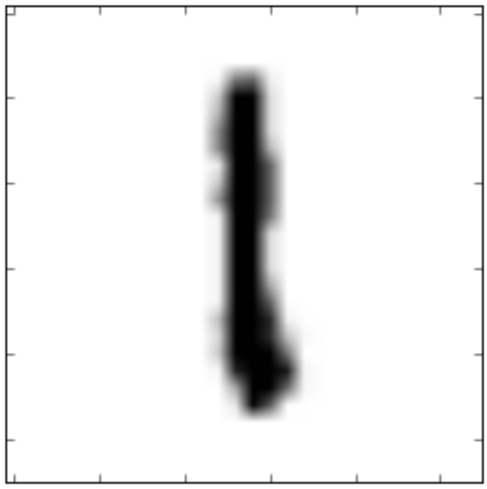
In [6]:
```
# Display several images in a row
def show(imgs, n=1):
    fig = pylab.figure()
    for i in range(0, n):
        fig.add_subplot(1, n, i+1, xticklabels=[], yticklabels=[])
        if n == 1:
            img = imgs
        else:
            img = imgs[i]
        pylab.imshow(img.reshape(28, 28), cmap="Greys")
```

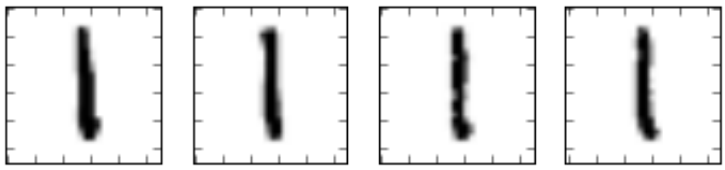
In [7]:
```
show(query_img)
show(train[result[0],:], len(result[0]))
```



In [8]:
```
########   ------------------QUESTION 2 :-----------------------
########  The other images in the dataset seem to perform with some er
ror. As can
########  be seen with the case of number 4. The number 4 is confused
with number
########  9 in this case which is likely caused due to prediction erro
r in the model.
########  Also, as will be seen in the further a few of the digits lik
e 1 and 7, 2
########  7 etc are also after confused.


query_img = test[1000]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

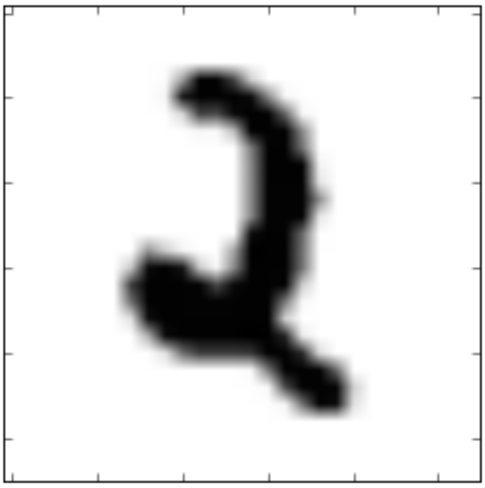
```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```

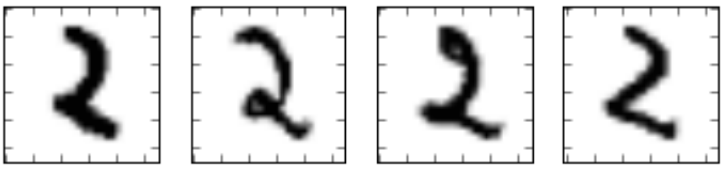




In [9]:
```
query_img = test[3000]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

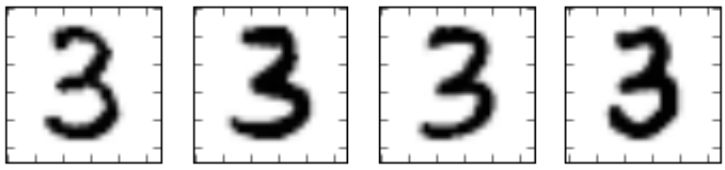
```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```





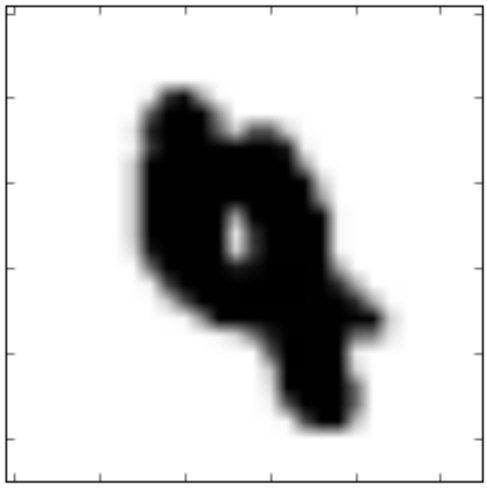
In [10]:
```
query_img = test[4000]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

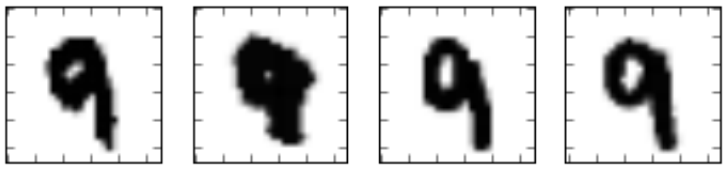
```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```



In [11]:
```
query_img = test[5000]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

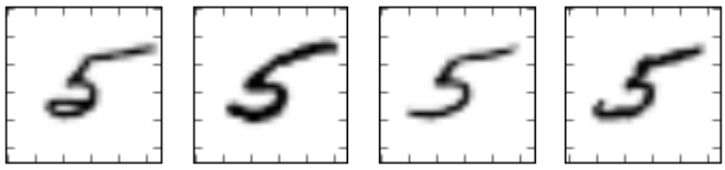
```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```





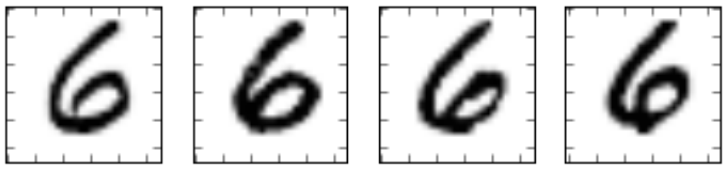
In [12]:
```
query_img = test[6000]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```



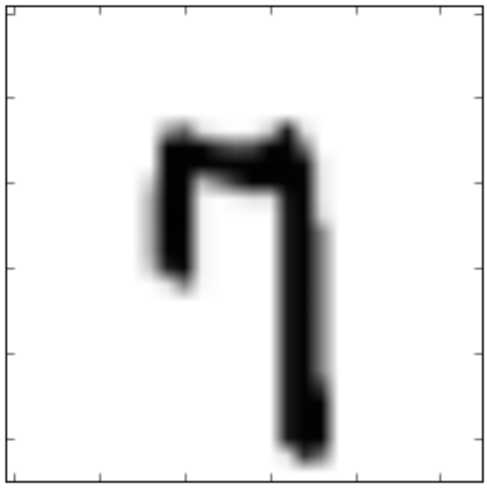
In [13]:
```
query_img = test[6500]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

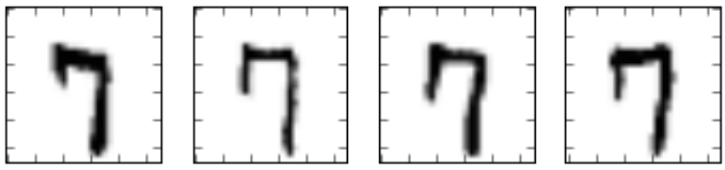
```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```



In [14]:
```
query_img = test[7000]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```



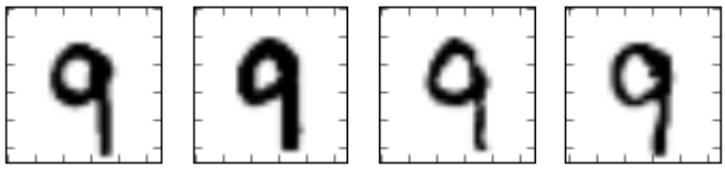


In [15]:
```
query_img = test[8200]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```

In [16]:
```
query_img = test[9200]
_, result = model.kneighbors(query_img, n_neighbors=4)
show(query_img)
show(train[result[0],:], len(result[0]))
```

```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```



In [17]:
```
train_labels = mnist.target[:60000]
test_labels = mnist.target[60000:]
test_labels_sample = test_labels[::50]
```

In [18]:
```
%%time
from sklearn.neighbors import KNeighborsClassifier
model = KNeighborsClassifier(n_neighbors=4, algorithm='brute').fit(tra
in, train_labels)
```

```
CPU times: user 7.27 ms, sys: 1.76 ms, total: 9.03 ms
Wall time: 7.74 ms
```

In [19]:
```
%%time
model.score(test_sample, test_labels_sample)
```

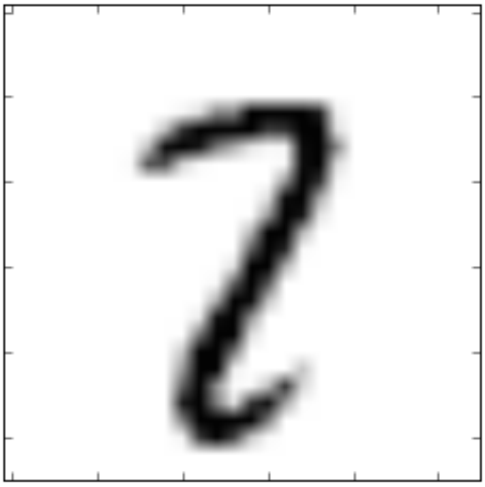
```
CPU times: user 1.15 s, sys: 247 ms, total: 1.4 s
Wall time: 870 ms
```

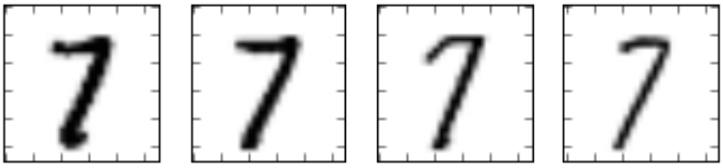
Out[19]:
```
0.95999999999999996
```

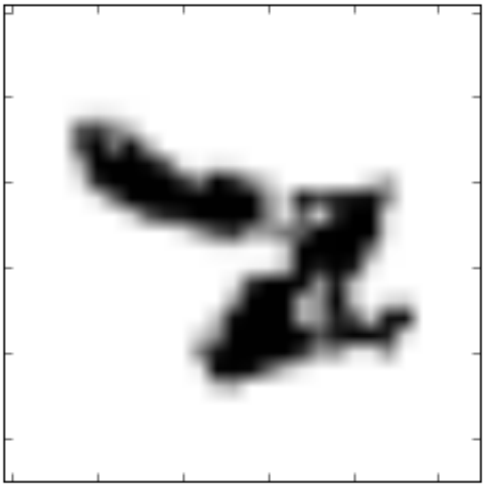
In [20]:
```
##########-------------Question 2 continued...
########## As can be seen below....There are digits which have been co
nfused
#########  These includes 4 and 9, 5 and 3, 5 and 8, 8 and 1, 7 and 1
etc.
#########

preds = model.predict(test_sample)
print (test_sample.shape)
errors = [i for i in range(0, len(test_sample)) if preds[i] != test_la
bels_sample[i]]

for i in errors:
    query_img = test_sample[i]
    _, result = model.kneighbors(query_img, n_neighbors=4)
    show(query_img)
    show(train[result[0],:], len(result[0]))
```

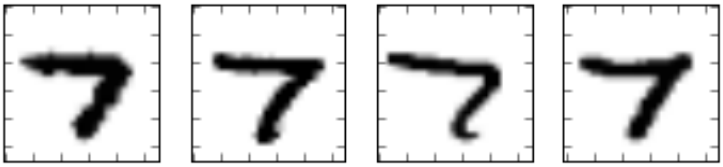
```
(200, 784)
```

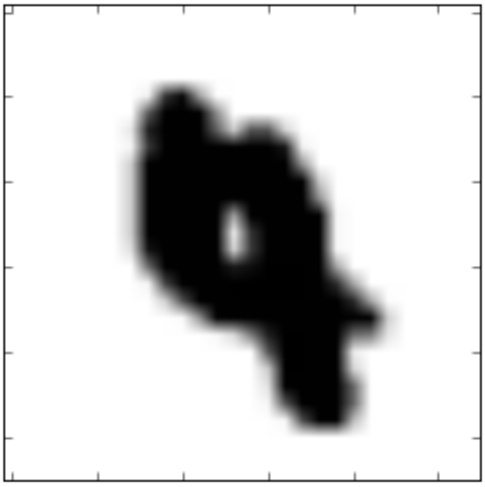
```
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
/Users/abhisheknigam/anaconda/lib/python3.5/site-packages/sklearn/ut
ils/validation.py:386: DeprecationWarning: Passing 1d arrays as data
is deprecated in 0.17 and willraise ValueError in 0.19. Reshape your
data either using X.reshape(-1, 1) if your data has a single feature
or X.reshape(1, -1) if it contains a single sample.
  DeprecationWarning)
```

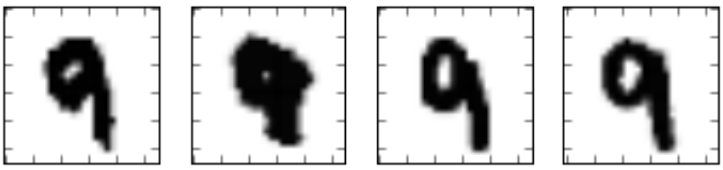


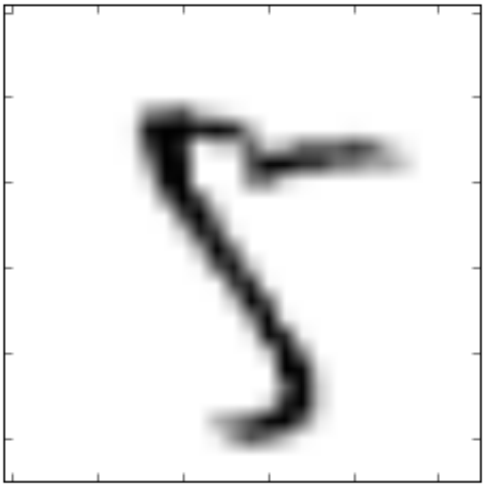


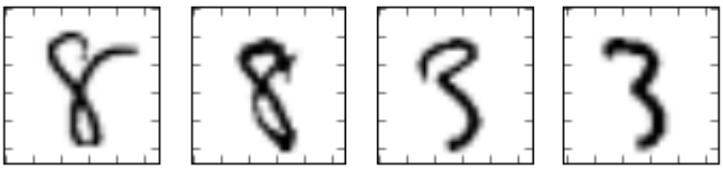


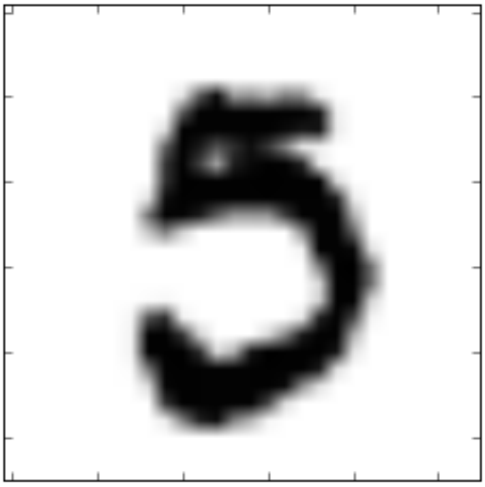


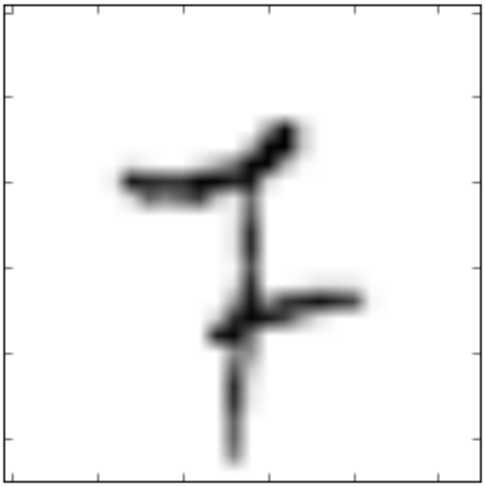


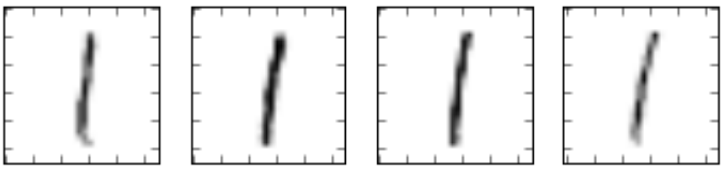


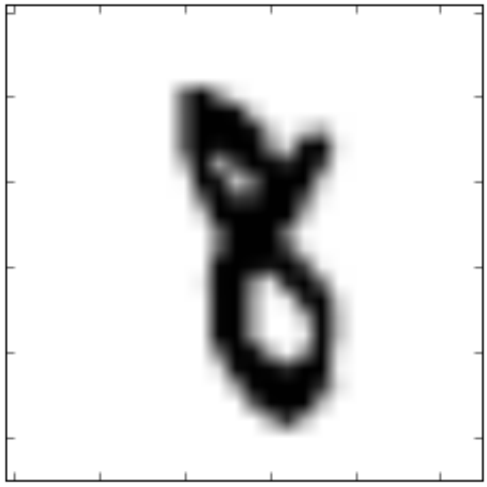


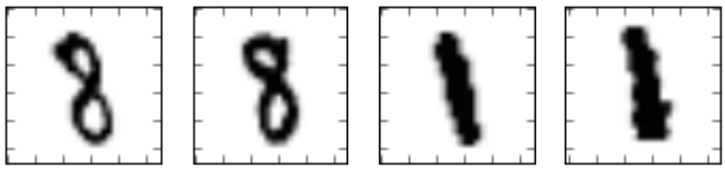


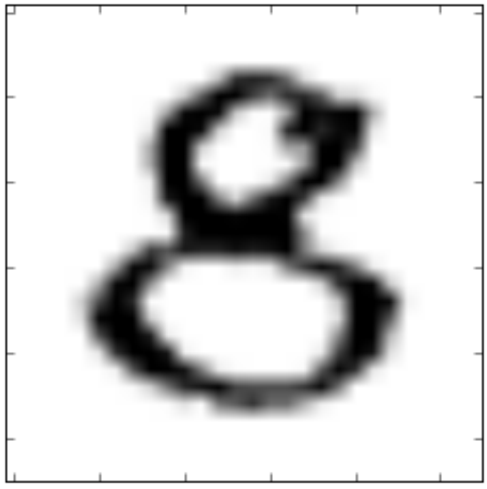


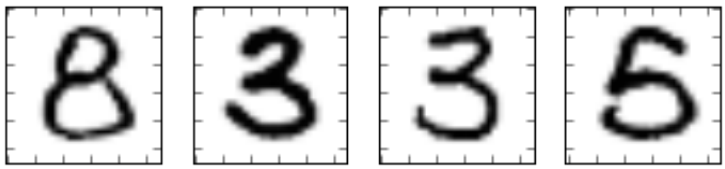












In [ ]: