In [39]:
```
import numpy as np
import pandas as pd
from sklearn import tree,cross_validation
cars=pd.read_csv('cars.csv.txt')
cars=cars.replace(['vhigh','high','med','low','5more','more','small','big','unacc
list=['buying','maint','doors','persons','lug_boot','safety']
X=cars[list]
y=cars['class ']
X=np.array(X)
y=np.array(y)
X_train,X_test,y_train,y_test=cross_validation.train_test_split(X,y,test_size=0.3
clf=tree.DecisionTreeClassifier()
clf.fit(X_train,y_train)
accuracy=clf.score(X_test,y_test)
print(accuracy)
```

```
0.976878612717
```

In [35]:
```
print(len(X_train))
print(len(y_train))
print(len(X_test))
print(len(y_test))
```

```
1209
1209
519
519
```

In [ ]: