In [1]:
```
import pandas as pd
```

In [2]:
```
data = pd.read_csv("data_lab_1.csv")
```

In [3]:
```
data.head(10)
```

Out[3]:

	sky	airtemp	humidity	wind	water	forecast	enjoysport
0	sunny	warm	normal	strong	warm	same	yes
1	sunny	warm	high	strong	warm	same	yes
2	rainy	cold	high	strong	warm	change	no
3	sunny	warm	high	strong	cool	change	yes
4	rainy	cold	high	strong	cool	change	no
5	sunny	warm	normal	strong	cool	same	yes

In [4]:

```
from functools import reduce

def find_dataset_info(data):
    ncol = data.shape[1] - 1
    uniqueValues = list(map(lambda x: len(data[x].unique()),data))[:-1]
    nInstances = reduce((lambda x, y: x * y), uniqueValues)

    uniqueValues = list(map(lambda x: x+2, uniqueValues))
    print(uniqueValues)
    nSyntactial = reduce((lambda x, y: x * y), uniqueValues)

    uniqueValues = list(map(lambda x: x-1, uniqueValues))
    nSemantic = 1 + reduce((lambda x, y: (x) * (y)), uniqueValues)

    print("Number of Instances = " + str(nInstances))
    print("Number of Syntactial = " + str(nSyntactial))
    print("Number of Semantic = " + str(nSemantic))

    return [nInstances, nSyntactial, nSemantic]

def list_to_string(ln):
    temp = ""
    for x in ln:
        temp+=x
        temp+=" "
    return temp

def findS_algorithm(data, positive_value = "yes"):
    hypothesis = ["Φ" for _ in range(data.shape[1]-1)]
    columns = data.columns
    print("Initial value of hypothesis: "+ list_to_string(hypothesis))
    first_value = True
    i = 0
    while(i < data.shape[0]):
        if(data.iloc[i,-1] != positive_value):
            i+=1
            continue
        value = list(data.iloc[i,:-1])
        if(first_value):
            hypothesis = list(data.iloc[i,:-1])
            first_value = False
        else:
            for item in range(len(value)):
                if(value[item] != hypothesis[item]):
                    hypothesis[item] = "?"
        print("Hypothesis value after data example "+ str(i+1)+": "+list
_to_string(hypothesis))
        i+=1
    return hypothesis
```

In [7]:
```
find_dataset_info(data)
```

```
[4, 4, 4, 3, 4, 4]
Number of Instances = 32
Number of Syntactial = 3072
Number of Semantic = 487
```

Out[7]:
```
[32, 3072, 487]
```

In [8]:
```
findS_algorithm(data)
```

```
Initial value of hypothesis: Φ Φ Φ Φ Φ Φ
Hypothesis value after data example 1: sunny warm normal strong warm sa
me
Hypothesis value after data example 2: sunny warm ? strong warm same
Hypothesis value after data example 4: sunny warm ? strong ? ?
Hypothesis value after data example 6: sunny warm ? strong ? ?
```

Out[8]:
```
['sunny', 'warm', '?', 'strong', '?', '?']
```