3. Assignment on Find S algorithm. Let's assume we have a dataset of Enjoysport with sevenn attributes: Sky, Air, Temp, Humidity, Wind, Water, Forecast, EnjoySport. Divide the Dataset into two groups: "Specific Hypothesis" and "Generic Hypothesis" using the Find-S algorithm.

In [1]:
```
#This dataset consists of seven attributes including the output.
#Let's import the required libraries.
import pandas as pd
import numpy as np
```

In [2]:
```
#Let us understand how to read the data of the CSV file(dataset).
dataset=pd.read_csv("ENJOYSPORT.csv")
```

In [3]:
```
dataset
```

Out[3]:

	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
0	Sunny	Warm	Normal	Strong	Warm	Same	1
1	Sunny	Warm	High	Strong	Warm	Same	1
2	Rainy	Cold	High	Strong	Warm	Change	0
3	Sunny	Warm	High	Strong	Cool	Change	1

The output of the above code would be the dataset EnjoySport.

Now, the next step is making an array of all attributes by excluding the output column.

In [4]:
```
arr=np.array(dataset)[:,:-1]
```

In [5]:
```
print(arr)
```

```
[['Sunny' 'Warm' 'Normal' 'Strong' 'Warm' 'Same']
 ['Sunny' 'Warm' 'High' 'Strong' 'Warm' 'Same']
 ['Rainy' 'Cold' 'High' 'Strong' 'Warm' 'Change']
 ['Sunny' 'Warm' 'High' 'Strong' 'Cool' 'Change']]
```

The next step is getting only the output values of the dataset.

In [6]:
```
target=np.array(dataset)[:,-1]
```

In [7]:
```
print(target)
```

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

- Instantiate the variable specific_hypothesis by the first positive example.
- Then for every positive example compare it with specific_hypothesis.
- If an attribute does not match, replace it with '?' else continue the process until the last positive example.

In [8]:
```
def train(Attributes,Target):

    for i, val in enumerate(Target):
        if val == 1:
```

```
            specific_hypothesis = Attributes[i].copy()
            break

    for i, val in enumerate(Attributes):
        if Target[i] == 1:
            for x in range(len(specific_hypothesis)):
                if val[x] != specific_hypothesis[x]:
                    specific_hypothesis[x] = '?'
                else:
                    pass

    return specific_hypothesis
```

The final value in specific_hypothesis is the most specific hypothesis of the dataset.

In [9]:
```
output=train(arr,target)
print("Hypothesis Space for Enjoy Sport Dataset",output)
```

```
Hypothesis Space for Enjoy Sport Dataset ['Sunny' 'Warm' '?' 'Strong' '?' '?']
```

This means that if the first four attributes of record are Sunny, Warm, High, Strong respectively then the output of that record is positive(Yes) irrespective of the last two attributes Water and forecast.

In [ ]:
```

```