# Notes - Algorithms & Data Structures - Sets

- Dr Nick Hayward

A brief intro to *sets*, and their general usage.

## Contents

- Intro
- Worked example
- Code example

## Intro

A set is an abstract data type, which stores unique values.

There is no pre-defined, discernible order to the data stored but the data must be unique.

Such data is a working implementation of a mathematical *finite set*.

Unlike many other data structures, we do not customarily retrieve a specific element from the set.

Instead, we check the set for the existence of a given element. It's unique record may then be used to retrieve the required data.
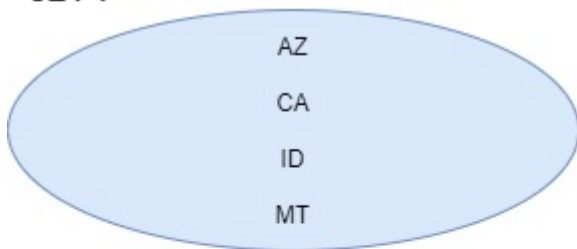
## Worked example

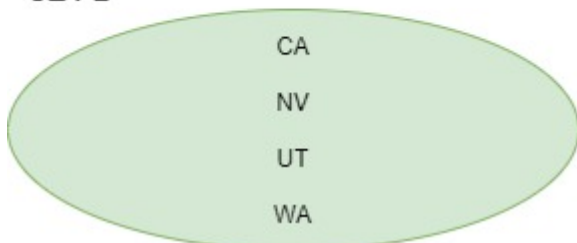We might represent sets of items, which will be unique to each set.

There may be duplication of elements in multiple sets, but the values must be unique per set.
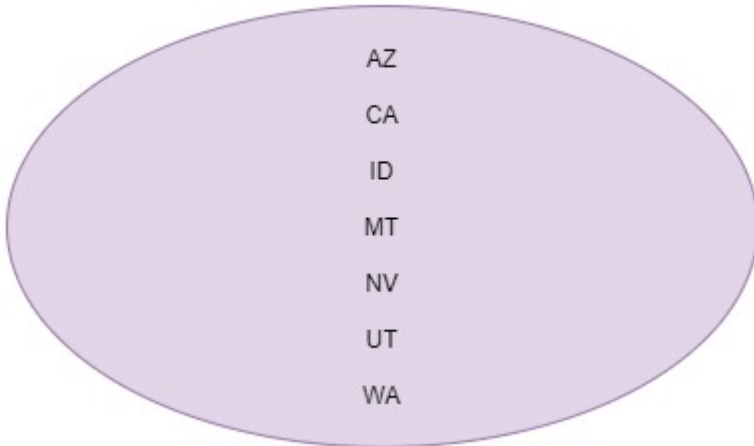
For example,

**SET 1**

AZ

CA

ID

MT

**SET 2**

CA

NV

UT

WA

We may then use such sets to perform various operations.

- union - a set containing all unique elements from a group of sets
  - combine sets to create a single unified set
  - e.g. union of set 1 and set 2
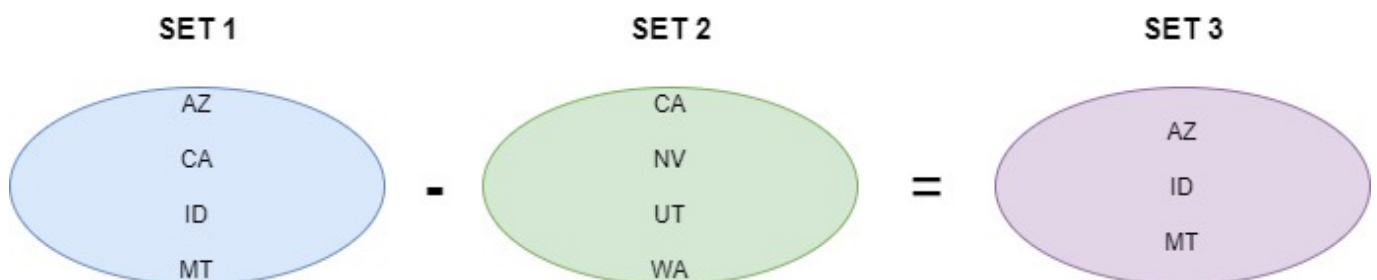
**UNION**

AZ

CA

ID

MT

NV

UT

WA

- intersection - elements that exist in each of the intersected sets
  - find the elements that exist into all of the defined sets
  - e.g. states that are in set 1 and set 2

**INTERSECTION**

AZ  MT  NV

CA  UT

ID  WA
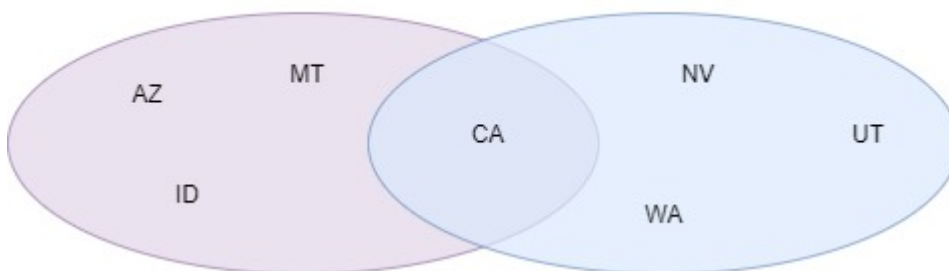
- difference - calculate the difference between defined sets
  - subtract elements in one set from elements in another set
  - e.g. subract elements in set 1 from elements in set 2

**DIFFERENCE**

| SET 1 | | SET 2 | | SET 3 |
|---|---|---|---|---|
| AZ | | CA | | AZ |
| CA | - | NV | = | ID |
| ID | | UT | | MT |
| MT | | WA | | |

**Code example**

We may implement such operations in code.

For example, in Python we may use a set as follows

```python
states_set1 = set(["az", "ca", "id", "mt"])
states_set2 = set(["ca", "nv", "ut", "wa"])

# set union
states_union = states_set1 | states_set2
# set intersection
states_intersect = states_set1 & states_set2
# set difference
states_diff = states_set1 - states_set2
```

If we then check the results of these operations, we may see the following output.

For example,

- union of sets

```
{'ut', 'nv', 'az', 'mt', 'id', 'wa', 'ca'}
```

- intersection of sets

```
{'ca'}
```

- difference of sets

```
{'az', 'id', 'mt'}
```