Sets: Properties and Operations

What is a Set?

A set is a collection of discrete data items. The members of the set can be numbers or names.

Describing a Set

There are two distinct ways of describing the members of a set. One is to list them explicitly, like you would find in a database of names.

A = { Mark, Angela, Frank, Laura }

A couple features of sets is that order doesn't matter, and duplicates don't really count.

{ Mark, Angela, Frank, Laura } = { Laura, Frank, Mark, Angela }

and

{ Mark, Angela, Frank, Laura } = { Mark, Laura, Angela, Mark, Frank, Laura }

One way to think of this is a database query using the "unique" keyword. Similarly, order is a matter of how *we want* the set contents listed. A class list can be ordered by student name, or student number, depending on what it is used for. We commonly remove duplicates for convenience, although they are occasionally used depending on the specific application.

Another way to define a set is to describe a mathematical relationship.

$A = \{x \mid 2x + 6 = 0 \}$

The vertical bar can be read as "such that", so that the entire statement would be read as

"set A consists of members solving for x, such that 2 times x plus 6 equals 0"

This same set can be listed explicitly.

$A = \{ -3 \}$

Variables used to identify sets are usually denoted with uppercase letters, such as A. Lowercase letters usually refer to specific member of a set by the uppercase name. Thus, a is the variable we would use to refer to a member of set A.

We can write this like this:

$a \in A$

This can read as "*a* is a member of set *A*", or "*a* is an element in set *A*". The symbol for the member can be any variable.

$x \in A$ (member / element of)

Similarly, we use a slightly different symbol to state that the content of a variable is not a member of a particular set.

$x \notin A$ (not a member / element of)

This notation is good for individual members, but what if we are trying to compare a group of set members? For that we have "*subsets*". A subset is any set whose members are members of another set.

A = { Mark, Angela }
B = { Mark, Angela, Frank, Laura }

Set A is a subset of set B because all members of set A are in set B. A symbol that is commonly used is $\subseteq$. Thus, we could write

$A \subseteq B$ (subset)

We make one additional distinction between sets, and that has to do with whether every member is accounted for. If every member is accounted for, the sets are equal. If they are not, we have a *proper subset*. A proper subset is denoted using a slightly different symbol.

$A \subset B$ (proper subset)

Thus, if two sets are the same, then one *cannot* be a proper subset of the other.

As for using the symbol, consider the way the symbol for *less than or equal to* is used compared to simply *less than*. '≤' means they may be equal, whereas '<' means that they may *not* be equal.

An additional feature of sets and subsets is that we can identify them as sets within sets, like this:

A = { {Mark, Angela }, { Frank, Laura} }

Which is the same as:

A = { Mark, Angela, Frank, Laura }

The number of members in a set can be determined using vertical brackets similar to "absolute value". Thus,

A = { Mark, Angela, Frank, Laura }

We can say that

$|A| = 4$

A set with no members is called an *empty set*, or *null set*. Two ways to describe this is

$\varnothing$

or simply

{ }

A *power set* is a collection (set) of sets which represents every valid subset of a set. The symbol for the power set is a stylized P, or $\mathscr{P}$**.** Thus, where we have a set…

B = { Fred, Mary, Jane }

The members of the power set for set B would be

$\varnothing$, {Fred}, {Mary}, {Jane}, {Fred, Mary},
{Fred, Jane}, {Mary, Jane}, {Fred, Mary, Jane}

We could also write

$\mathscr{P}$ (B) = { $\varnothing$, {Fred}, {Mary}, {Jane}, {Fred, Mary},
{Fred, Jane}, {Mary, Jane}, {Fred, Mary, Jane} }

All of these are proper subsets of set B except for {Fred, Mary, Jane}. Also notice that we covered each combination of each set member where order doesn't matter.

The number of members of the power set would be written as

$|\mathscr{P}(B)|$

Notice that in the case above the number of elements in set B was 3. The number of elements in the power set of B is 8. If you were to create power sets for other sized sets you would see that there is relationship between the two. That relationship can be described here:

Where

$|B| = 3$,

$|\mathcal{P}(B)| = 8 = 2^3$

$|\mathcal{P}(B)| = 2^{|B|}$

Set Operations

Operations between sets allow us to examine and manipulate the contents of sets in ways similar to logical and Boolean operations. For the following examples, we will define two sets, *A* and *B*.

A = { Mary, Mark, Fred, Angela, Frank, Laura }
B = { Fred, Mary, Frank, Jane }

Union The *union* of two sets is a new set which combines all of the members of both sets (and discards duplicates). Thus, the union of sets *A* and *B* would look like this:

$A \cup B$ = { Mary, Mark, Fred, Angela, Frank, Laura, Jane }

One way to remember the symbol for set union is that it looks like an uppercase 'U'.

Intersection The *intersection* of two sets is a new set which only includes those members present in both sets. Thus, the intersection of sets *A* and *B* would look like this:

$A \cap B$ = { Mary, Fred, Frank }

Two sets are said to be ***disjoint*** if their intersection is empty. Thus, sets *X* and *Y* are disjoint if

$X \cap Y = \varnothing$ (disjoint)

Subtraction Subtraction, *per se*, does not actually exist in set operations, but we can perform an operation that looks very similar. The operation $A - B$ removes those members in set *B* that are in set *A*. If a member in set *B* isn't in set *A*, then nothing is done. (There is no sense of "negative data", so you cannot remove what isn't there.) In our case we would get this:

A = { Mary, Mark, Fred, Angela, Frank, Laura }
B = { Fred, Mary, Frank, Jane }

A – B = { Mark, Angela, Laura }

Universal Set / Universe The *universal set* is defined as a set with every element of our data. You could think of it as the union of all subsets. Another way is to think of the universal set as the entire database. The universal set, or *universe*, is usually identified with a capital U.

In our case, the universal set would be:

U = { Mary, Mark, Fred, Angela, Frank, Laura, Jane }

Complement The *complement* of a set is defined as every member in the universal set that is not in the given set. This is identified by using a bar over the set name, for example: $\overline{X}$.

U = { Mary, Mark, Fred, Angela, Frank, Laura, Jane }
C = { Mary, Mark, Angela, Frank, Laura }

$\overline{C}$ = { Fred, Jane }

Another way to see this is the relationship:

$U - C = \overline{C}$

Logical Properties

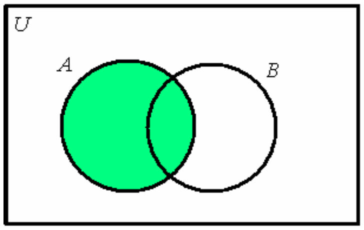
Like propositional and Boolean operators there are a set of properties which can be applied to sets.

Associative Law	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
Commutative Law	$A \cup B = B \cup A$ $A \cap B = B \cap A$
Distributive Law	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Identity Law	$A \cup \varnothing = A$ $A \cap U = A$
Complement Law	$A \cup \overline{A} = U$ $A \cap \overline{A} = \varnothing$
Idempotent Law	$A \cup A = A$ $A \cap A = A$
Bound Law	$A \cup U = U$ $A \cap \varnothing = \varnothing$
Absorption Law	$A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$
Involution Law	$\overline{\overline{A}} = A$
0/1 Law	$\overline{\varnothing} = U$ $\overline{U} = \varnothing$
DeMorgan's Law	$\overline{\mathrm{A} \cup B} = \overline{A} \cap \overline{B}$ $\overline{\mathrm{A} \cap \mathrm{B}} = \overline{A} \cup \overline{B}$

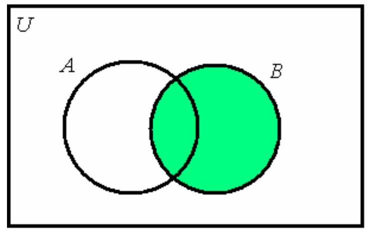
Venn Diagrams

Venn diagrams are a method for displaying set relations in a way that is more visual than the more algebraic form used above. A circle or "blob" shape is used to represent each subset. The portions of the diagram that are shaded represent the set expression.

$A =$

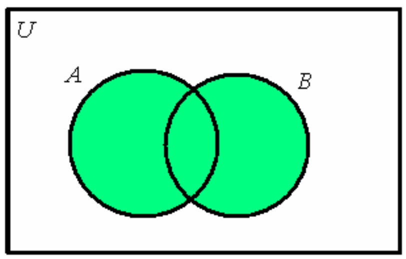


$B =$



Union The union of two or more sets can be shown in a diagram as shown here.

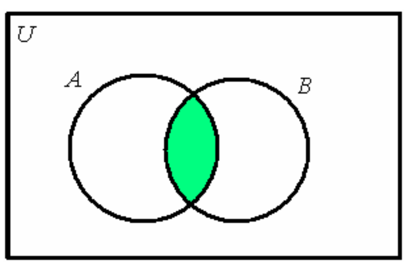
$A \cup B =$



Notice that the contents of both circles are shaded.

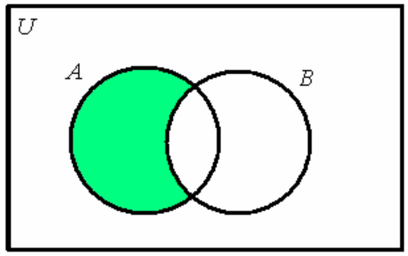
Intersection The intersection of two or more sets can be shown in a diagram as shown here.

$A \cap B =$



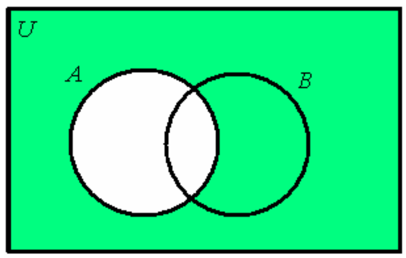
Notice that only the portion shared by both circles is shaded.

Subtraction The subtraction of two sets can be shown in a diagram as shown here.
$A - B =$



Notice that the contents of the circle representing set *A* is shaded but not shaded where there is an overlap with the circle representing set *B*.

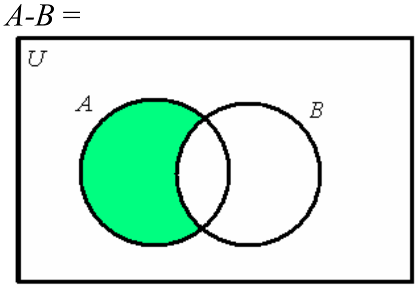
Complement The complement of a set is everything in the universe that is not in the set. The subtraction of two sets can be shown in a diagram as shown here.



Notice that everything outside of set A is shaded.

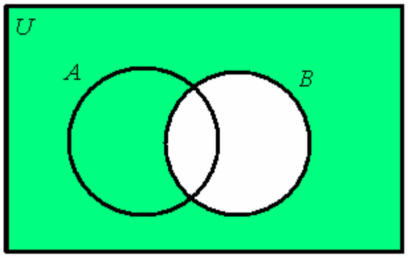
One thing that can be done with Venn diagrams is to help visualize a process. For example, the operation *A-B* may require an interpretation in order to become more useful (because subtraction cannot be manipulated like the other operators). What set of operations is equivalent to *A-B* ?

$A\text{-}B =$

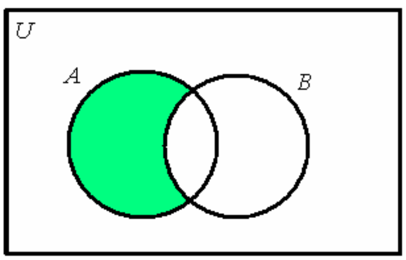


The complement of set *B* gets us this:

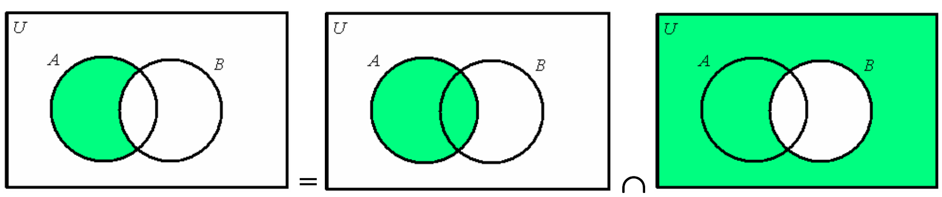
$\overline{B}$ =



Intersecting this with set *A* gives us this:
(intersection shades the portions in common)



Thus, we could reasonably say that $A\text{-}B = A \cap \overline{B}$.



Set Relations

Relations between sets consist of little more than linking data elements in a way that makes them more useful to us. A simple example is a student's name along with their student ID number.

(Mark Smith, 65297)

Also called an ordered pair, this type of data is commonly represented as a sequence of data in parentheses. The link is commonly referred to as a tuple (no, that is not a misspelling). If more than two data elements are associated, then it is referred to as an *n*-tuple, where *n* is the number of elements in the ordered list. Another name may be a record in a non-relational database.

If we have a set of ordered pairs…

(Fred, 832)
(Mary, 719)
(Frank, 106)
(Jane, 521)

They can be put in the form of a table like this:

Fred	832
Mary	719
Frank	106
Jane	521

Ordered pairs may be represented in the form of a set, but the pairing must be preserved.

R = { (Fred, 832), (Mary, 719), (Frank, 106), (Jane, 521) }

It may be helpful to think of this as a database, where specific order of the pairs doesn't matter, but each pair represents its' own record in the database.

Functions

A set of ordered pairs or tuples establishes an explicit link between the elements of the ordered pair/list.

Let set X be a set of students.

X = { Mark, Angela, Frank, Laura }

Let set Y be a set of classes.

Y = { CIS101, MAT114, MAT131, CIS310, CIS121, RHT103 }

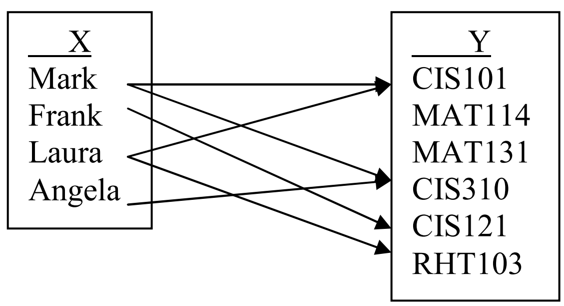
A set of relations may be defined like this:

R = { (Mark, CIS101), (Frank, CIS121), (Laura, CIS101), (Mark, CIS310), (Angela, CIS310), (Laura, RHT103) }

We can write that if (Frank, CIS121) $\in R$, then we can also write

Frank R CIS121

When placed into table form…



We can use arrows to help define the relation between the left column of the table, and the right column. Notice that we are not obligated to use all elements in all sets.