Algorithms and Data Structures (CSci 115)

California State University Fresno
College of Science and Mathematics
Department of Computer Science
H. Cecotti

Learning Objectives

- **Definitions**
  - Notations
  - Subsets, Powersets
- **Operation on sets**
  - Union, Intersection, Difference, Complement
- **Concepts and implementation of**
  - Sets
    - Array of boolean
  - Multisets
    - Array of ints

Set definitions

- **A set is a group of "objects" e.g.**
  - People in a class: { "Anne", "Bill", "Colum" … }
  - Modules in a Course {"COM812", "COM814", "COM525" … }
  - Colors of a rainbow: { "red", "orange", "yellow", "green", … }
  - All positive numbers <= 5: {1, 2, 3, 4, 5}
  - A few selected real numbers: { 2.1, ∏, 0, -6.32, e }
  - Sets can contain non-related elements:

    { 3, 'a', "green", 12.4, false}
- **Sets are listed and surrounded with curly brackets (braces)**

Set Properties

- **Sets do not have duplicate elements**
  - ➢ Consider the set of vowels in the alphabet.
    - o It makes no sense to list them as

      {‘a’, ‘a’, ‘e’, ‘i', ‘o’, ‘o’, ‘o’, ‘u’}
    - o What we really want is just {‘a’, ‘e’, ‘i', ‘o’, ‘u’}
- **Order does not matter**
  - ➢ We often write them in order as it is easier for us to understand them in that way

    {1, 2, 3, 4, 5} is equivalent to {3, 5, 2, 4, 1}

Specifying a set…

- ***Sets*** are usually represented by a capital letter

  A, B, C …. S etc.

- ***Elements*** are usually represented by an *italic* lower-case letter

  a, *x*, *y*, etc.

- Easiest way to specify a set is to list all the elements:

  A = {*1*, *2*, *3*, *4*, *5*}

  - Not always possible for large or infinite sets

- We can use an ellipsis (…)
- This allows us to specify (say) etcetera.
- EXAMPLES:

  Set B = {*0*, *1*, *2*, *3*, *…*} // an *infinite* set

  All the values greater than 0

  Set C = {*1*, *2*, *3*, *…* , *18*, *19*, *20*} // a *finite* set

  All the values between 1 and 20 (inclusive)

Specifying a set…

- Can use *set-builder* notation

D = {x | x is prime and $x > 2$}

E = {x | x is odd and $x > 4$}

The vertical bar should be read as: *"such that"*

Thus, **Set D** is read (in English) as

*"All elements **x such that x is prime** AND **x is greater than 2**"*

This **Set E** is read (in English) as:

*"All elements **x such that x is odd** AND **x is greater than 4**"*

Specifying a set

- A set is said to "***contain***" the various "*members*" or "*elements*" that make up the set

- If an element ***a*** is a member of (or is an element of) set S, we use then notation $\boldsymbol{a} \in \mathbf{S}$
  - Hence: $\mathbf{4} \in \{\mathbf{1, 2, 3, 4}\}$

- If an element is not a member of (or is an element of) a set S, we use the notation $\boldsymbol{a} \notin \mathbf{S}$
  - Hence: $\mathbf{7} \notin \{\mathbf{1, 2, 3, 4}\}$

    $\mathbf{0} \notin \{\mathbf{1, 2, 3, 4}\}$

Some Commonly Used Sets

- **N** = {0, 1, 2, 3, …}
  - The set of ***natural numbers***
- **Z** = {…, -2, -1, 0, 1, 2, …}
  - The set of ***integers***
- $\mathbf{Z}^+$ = {1, 2, 3, …}
  - The set of ***positive integers***
- **Q** = $\{p/q \mid p \in \mathbf{Z},\ q \in \mathbf{Z},\ q \neq 0\}$
  - The set of **rational numbers**
  - Any number that can be expressed as a fraction of two integers (where the divisor (bottom integer) is not zero)
- **R** : the set of real numbers
- Ø is an empty or null set
  - Ø = {}
  - Ø is a subset of any set A
    - (see subset…)
  - Contains no elements
    - | Ø | = 0 (see cardinality…)

Universal sets

- ***U* is the universal set**
  - The set of ALL of elements (or the "universe") from which given any set is drawn
- **Examples**
  - For the set {-2, 0.4, 2} ***U*** would be the ***real numbers***
  - For the set {0, 1, 2}, ***U*** would/could be the ***natural numbers*** (integers from zero and upwards)
  - For the set of the ***students*** in this class, ***U*** would be all the ***students in the University*** (or all the ***students/people in Ireland***)
  - For the set of the ***vowels*** of the alphabet, ***U*** would be all the ***letters of the alphabet***

Upper and lower bounds

- **Definitions**
  - The least upper bound of a given set of real numbers =
    - the smallest number bounding this set from above; its greatest lower bound is the largest number bounding it from below.
  - Let there be given a subset X of the real numbers.
  - A number β = its **least upper bound** (sup X (supremum)), if $\forall\ x \in X$ satisfies the inequality $x \leq \beta$, and if for any $\beta' < \beta\ \ \exists\ \ x' \in X \mid x' > \beta'$.
  - A number α = the **greatest lower bound** of X (inf X (infimum)), if $\forall\ x \in X$ satisfies the inequality $x \geq \alpha$, and if $\forall\ \ \alpha' > \alpha\ \exists\ x' \in X \mid x' < \alpha'$

Upper and lower bounds

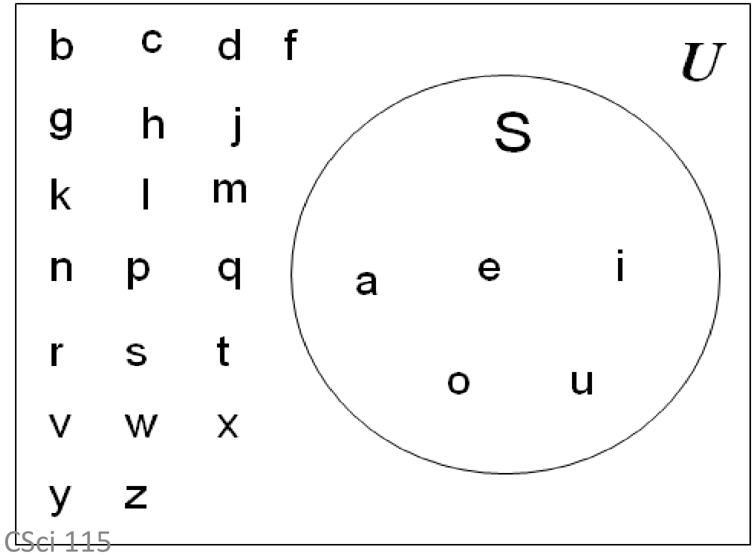
- **Examples**:
  - ➢inf(a,b)=a   sup(a,b)=b   inf(a,b)=a   sup(a,b)=b;
  - ➢If the set X consists of two points a and b, a<b, then inf X=a, sup X=b
    - →the least upper bound (greatest lower bound) may either belong to the set (case of the interval [a,b]) or not belong to it (case of the interval (a,b)).
    - →If a set has a largest (smallest) member, this number will clearly be the **least** upper bound (greatest lower bound) of the set.
- The least upper bound (greatest lower bound) of a set not bounded from above (from below) = symbol $+\infty$ (respectively, by the symbol $-\infty$)
  - o If N={1,2,…} is the set of natural numbers,
    - then Inf N=1 and sup N=$+\infty$.
  - o If Z is the set of all integers (positive and negative),
    - then inf Z=$-\infty$ and sup Z=$+\infty$.

Majorant and minorant

- **Definitions**
  - A majorant of a subset X of an ordered set E is
    - an element $y \in E$ | that $y \geq x$ for every $x \in X$.
  - A minorant of a subset X of an ordered set E is
    - an element $y \in E$ | that $x \geq y$ for every $x \in X$.

Venn diagrams - introduced by John Venn (1880)

- **Venn Diagrams** allow us to represent sets graphically
  - The **box** represents the universal set
  - **Circles** represent the set(s)
- Consider set **S**, which is the set of all vowels in the alphabet
- The individual elements are usually not written in a Venn diagram



Set Equality

- **Two sets are equal if they have the same elements**

  {1, 2, 3, 4, 5} = {5, 4, 3, 2, 1}

  - Recall: Order does not matter!

  {1, 2, 3, 2, 4, 3, 2, 1} should be {4, 3, 2, 1}

  - Recall: Duplicate elements do not matter!

- **Two sets are not equal if they do <u>not</u> have the same elements**

  {1, 2, 3, 4, 5} ≠ {1, 2, 3, 4}

Subsets

- Set S : subset of Set T if and only if all the elements of S are also elements of T
  - S = {2, 4, 6} and T = {1, 2, 3, 4, 5, 6, 7}
    - S is a subset of T
    - Notation: **S $\subseteq$ T**

    Or by {2, 4, 6} $\subseteq$ {1, 2, 3, 4, 5, 6, 7}
- If S is **not** a subset of T, it is written as such: S $\nsubseteq$ T
    - For example, {1, 2, ***8***} $\nsubseteq$ {1, 2, 3, 4, 5, 6, 7}
- Any set is a subset of itself!
  - Thus, for any set R, R $\nsubseteq$ R

Power Sets

- The Power Set of a set **A** is a set containing **all subsets** of **A**.
  - A set of n elements has 2^n subsets
    - **including both the set itself and the empty set**
  - The total number of distinct k-subsets on a set of n elements given by the binomial sum

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \quad \text{with} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

k=1
n=4
→ 4!/(1!*(4-1)!
= 4!/3! = 4

- Written as: A: $\mathcal{P}$(A)
- Examples:

If A = {1} $\mathcal{P}$**(A)** = { {}, {1} }

If B = {1, 2} $\mathcal{P}$***(B)*** = {∅, {1}, {2}, {1, 2}}

If C = {1, 2, 3} $\mathcal{P}$***(C)*** = {{}, {1}, {2}, {3}, {1, 2}, {1, 3}, {2, 3}, {1, 2, 3} }

If D = {a, b} $\mathcal{P}$***(D)*** = {{}, {a}, {b}, {a, b}}

Midterm & final: add the empty set in the list of sets, even if its cardinality is 0.

Proper Subsets

- If **A is a subset of B**
  - then A is also a ***proper subset of*** **B** as long as **A** is not equal to **B**
- A proper subset is written as **S** $\subset$ **T**
- A proper subset S_2 of a set S_1 is a subset that is **strictly** contained in S and so necessarily excludes at least one member of S_1.
  - The empty set is therefore a proper subset of any nonempty set.

Proper Subsets

(1) **S** is a proper subset of **T**

S is not equal to T, and S is a subset of T

A proper subset is written as S $\subset$ T

(2) **R** is a subset of **T** BUT as **R** is equal to **T**

R is **not** a proper subset of **T**

Can be written as: R $\subseteq$ T and R $\not\subset$ T (or just R = T)

(3) **Q** is not a subset of **T** and thus is not a proper subset of **T**

T = {0, 1, 2, 3, 4, 5}   **S** = {1, 2, 3},

R = {0, 1, 2, 3, 4, 5}.   **Q** = {4, 5, 6}.

Set Cardinality

- **The *cardinality* of a set is the number of elements in a set**
- Written as **| A |**
- Examples

If R = {1, 2, 3, 4, 5} then **| R | = 5**

|∅| = 0, |{}|=0; empty set ∅ ={}

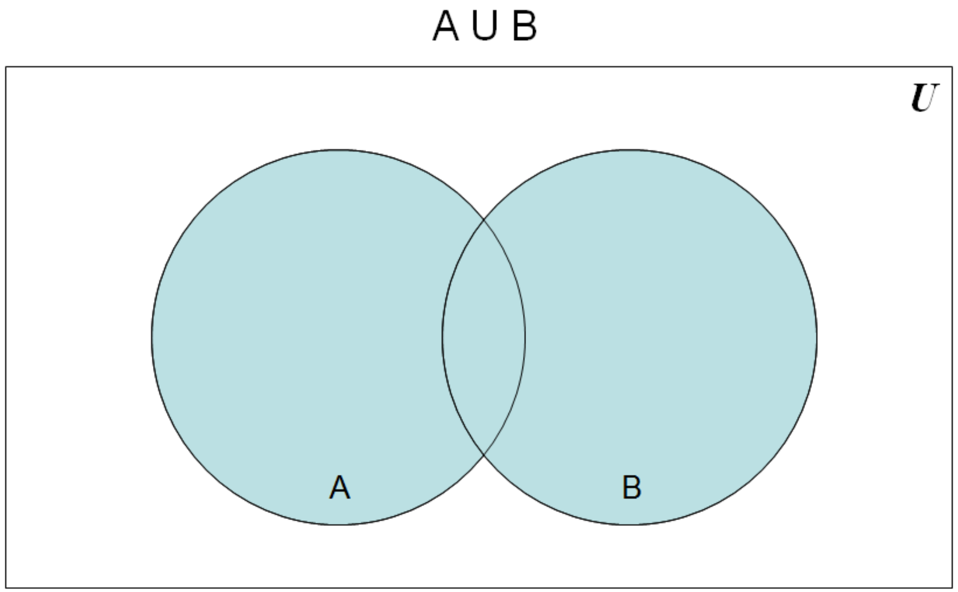
Let S = { ∅, {a}, {b}, {a, b} }. **| S | = 4**

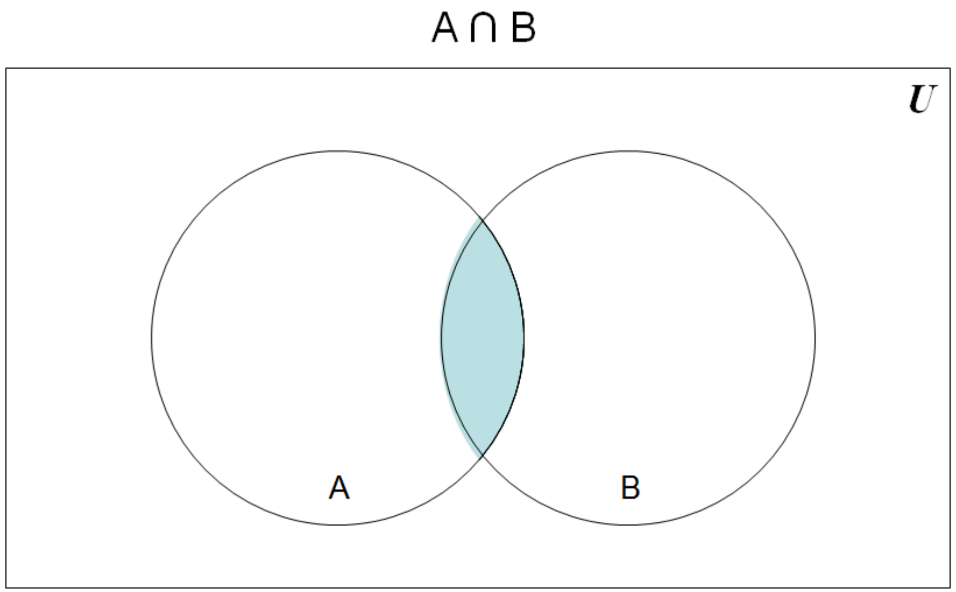
Let W = {1, 2, 3, {4, 5, 6}, 7, 8} **| W | = 6**

Set operations:

- Union

Intersection





Set operations: Union

- A Union of two sets A and B contains all of the elements in set A plus all the elements in set B
- Formal definition for the union of two sets:

  A U B = { x | $x \in$ A **or** $x \in$ B }

- Also:
  - {1, 2, 3} U {3, 4, 5} = {1, 2, 3, 4, 5}
  - {Clovis, Madera} U {3, 4} = {Clovis, Madera, 3, 4}
  - {1, 2} U $\varnothing$ = {1, 2};

    //No elements in Ø

Set operations: Intersection…

- The intersection of two sets A and B, contains all of those elements that are present in BOTH sets A and B
- Formal definition for the intersection of two sets:

  $A \cap B = \{ x \mid x \in A \textbf{ and } x \in B \}$

- Further examples
  - $\{1, 2, 3\} \cap \{3, 4, 5\} = \{3\}$
  - $\{\text{Clovis,Fresno}\} \cap \{3, 4\} = \varnothing$ as no elements in common
  - $\{1, 2\} \cap \varnothing = \varnothing$
  - **Any set intersection with the empty set gives the empty set**

Examples: Union & Intersection

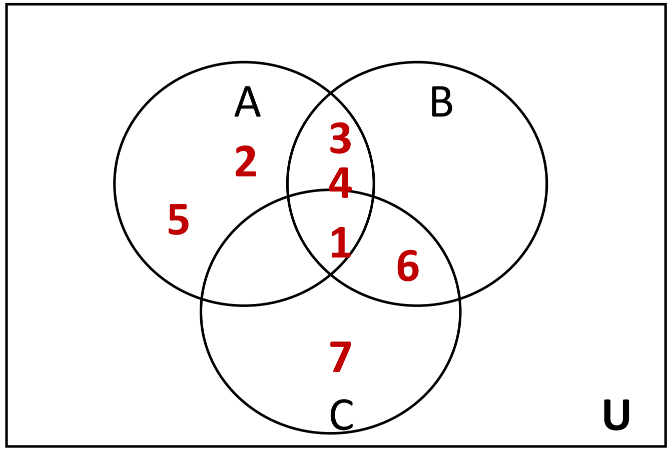
Let A = {1, 2, 3, 4, 5}, B = {1, 3, 4, 6}, and C = {1, 6, 7}

Find the following:

a. $B \cap C$ = {1, 6}

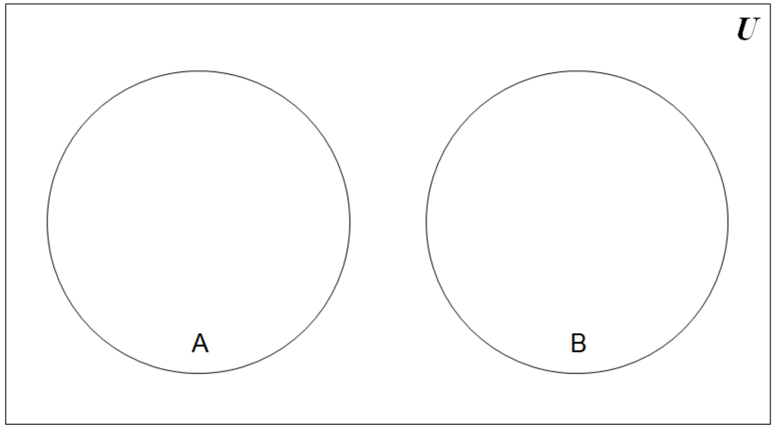
b. $A \cup B$ = {1, 2, 3, 4, 5, 6}

c. $A \cup (B \cap C)$ = {1, 2, 3, 4, 5, 6}



Disjoint Sets

- 2 sets are **disjoint**
  - if the have NO elements in common
  - if their intersection is the empty set
- Examples
  - {1, 2, **3**} and {**3**, 4, 5} are **not** disjoint
  - {“Goleta”, “Santa Barbara”} and {3, 4} are disjoint
  - {1, 2} and ∅ are disjoint
    - Their intersection is the empty set
  - ∅ and ∅ are disjoint!
    - Their intersection is the empty set



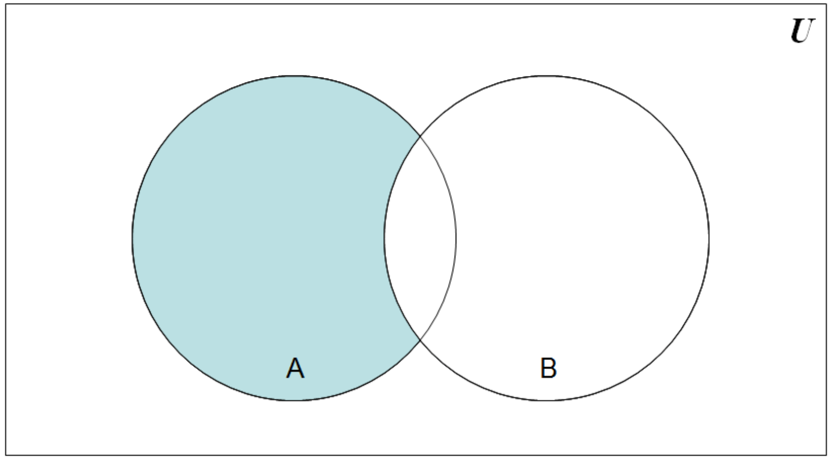
Set operations: Difference…

- The Difference of 2 sets corresponds to the elements in one set that are NOT in the other
- Formal definition for the difference of two sets:

  $A - B = \{ x \mid x \in A \textbf{ and } x \notin B \}$

- Further examples
  - $\{1, 2, \boldsymbol{3}\} - \{\boldsymbol{3}, 4, 5\} = \{1, 2\}$
  - {Clovis,Fresno} - {3, 4} = {Clovis,Fresno}
  - $\{1, 2\} - \varnothing = \{1, 2\}$
  - Difference of a set S with the empty set will be the set S

A - B



Warning

- Consider the following 2 sets:

  A = {1, 2, 3, 4 ,5}

  B = {0, 2, 4, 6, 8}

  A – B is simply {1, 3, 5}

  B – A is simply {0, 6, 8}

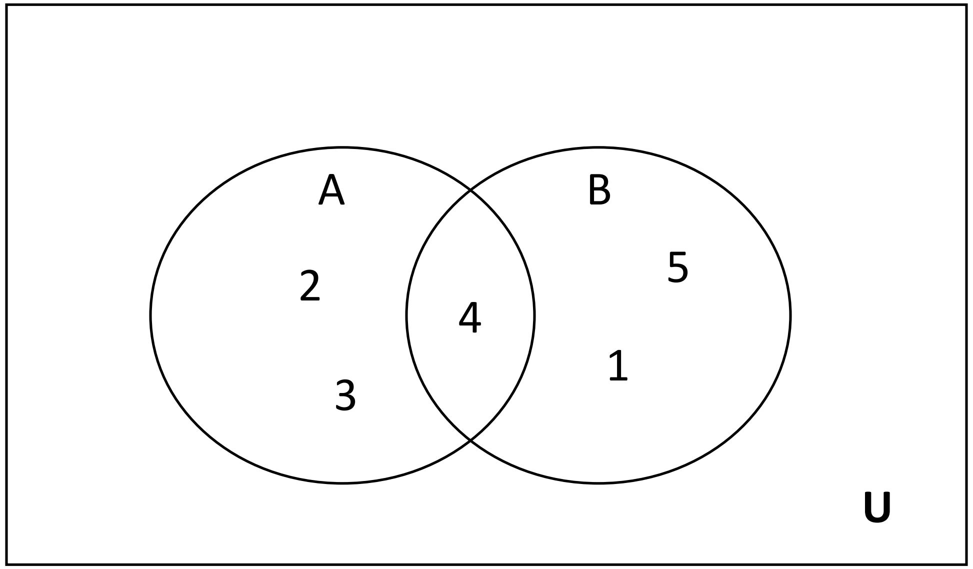
So clearly: (A – B) is NOT the same as (B – A)

Set operations: Difference

Let U = {1, 2, 3, 4, 5},

A = {2, 3, 4}

B = {1, 4, 5}



Find each specified set.

a. $U - B$ = {2, 3}

b. $B - A$ = {1, 5}

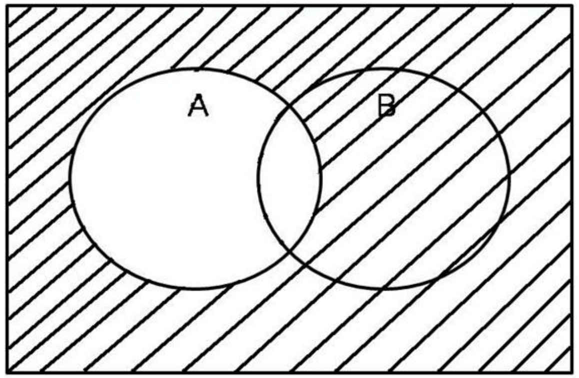
Set operations: Complement…

- Let *U* be the universal set, and let A be a subset of *U*.
- The **complement** of A, denoted by **A'**, is the set of elements in *U* that are not in A.
- That is

$$A' = \{x \mid x \in U \; and \; x \notin A\}$$

This set is also symbolized by

U – A .



Set operations: Complement

Let U = {a, b, c, d, e, f}

A = {a, c, e}

B = {b, d, e, f}

C = {a, b, d, f}.

Find each specified set.

a. $(A \cap B)'$ = {a, b, c, d, f}

b. $(A \cup B)'$ = ∅

c. $(A \cup B) \cap C'$ = {c, e}

d. $C \cup (A \cap B)'$ = {a, b, c, d, f}

Examples

Use the numbered regions of the diagram below to identify each specified set.

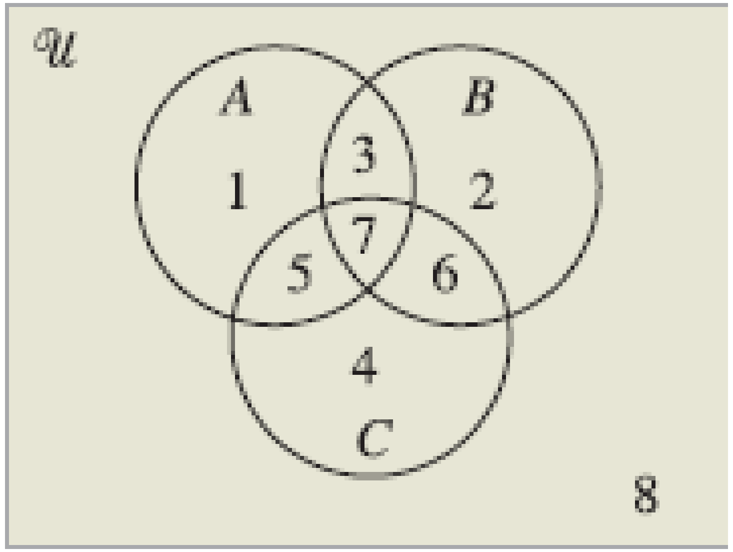
a. $A \cup B$ = {1, 2, 3, 5, 6, 7}

b. $B \cap C$ = {6, 7}

c. $A \cap B \cap C$ = {7}

d. B' = { 1, 4, 5, 8}

e. $(A \cup C) - B$ = {1, 4, 5}



Computer representation of sets

- Assume that U is finite (and reasonable!)
  - Let U be the alphabet
    - Each bit represents whether the element in U is in the set
- The vowels in the alphabet:

a b c d e f g h i j k l m n o p q r s t u v w x y z

1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0

- The consonants in the alphabet:

a b c d e f g h i j k l m n o p q r s t u v w x y z

0 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1

Computer representation of sets

- Consider the ***union*** of these two sets:

```
  1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0
∨ 0 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

- Consider the ***intersection*** of these two sets:

```
  1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0
∧ 0 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1
  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Computer representation of sets

- **Representing a Set of integers:**

  boolean [] setInt = new boolean [10];

  Index 0 to 9 represent integers,
  - Each element of the array is a boolean
  - That means they can be set to either true or false
  - Example: If the integer 6 (say) exists in setInt, the element at location 6 should be set as true

  {**0**, **2**, **6**, **7**} can be represented as {**t**, f, **t**, f, f, f, **t**, **t**, f, f}

  Where f means false, and t means true.

Set implementation: Set Union

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	f	f	t	f	f

S2	0	1	2	3	4	5	6	7	8	9
	t	t	f	f	t	f	f	t	t	f

Let S3 be the Union of S1 and S2, which logical operator should be used? OR : ||

S3[i] = S1[i] || S2[i];

Where i = 0, 1, …, 9;

S3	0	1	2	3	4	5	6	7	8	9
	t	t	f	t	t	f	f	t	t	f

Set implementation: Set Intersection

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	f	f	t	f	f

S2	0	1	2	3	4	5	6	7	8	9
	t	t	f	f	t	f	f	t	t	f

Let S3 be the Intersection of S1 and S2, which logical operator should be used? AND: &&
S3[i] = S1[i] && S2[i];
Where i = 0, 1, …, 9;

S3	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	f	f	t	f	f

Set implementation: Set Difference

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	f	f	t	f	f

S2	0	1	2	3	4	5	6	7	8	9
	t	t	f	f	t	f	f	t	t	f

Let S3 be the Difference of S1 and S2 (S1 – S2):

S3[i] = true (if S1[i] = true && S2[i]= false);

Where i = 0, 1, …, 9;

S3	0	1	2	3	4	5	6	7	8	9
	f	f	f	t	f	f	f	f	f	f

Set implementation: Add an Element

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	f	f	t	f	f

Add 5 to S1: changed

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	**t**	f	t	f	f

Add 7 to S1: no change

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	t	f	**t**	f	f

Set implementation: Remove an Element

-

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	f	f	t	f	f

remove 5 to S1: no change

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	**f**	f	t	f	f

remove 7 to S1: changed

S1	0	1	2	3	4	5	6	7	8	9
	t	f	f	t	f	t	f	**f**	f	f

Multisets

- **Multiset** (or **bag**) is a generalization of a *set*
- **Members are allowed to appear more than once**.
- In multisets the *order* of elements is irrelevant:
  - The multisets {*a*, *a*, *b*} and {*a*, *b*, *a*} are equal.
- Examples:
  - Set: {a, b}
  - Multiset: {a, b, a} {a, a, b, b, b, a}
  - Given a list of people (by name) and ages (in years), we could construct a multiset of ages, which simply counts the number of people of a given age.
- The number of times an element belongs to the multiset is the **multiplicity** of that member.
- The total number of elements in a multiset, including repeated memberships, is the **cardinality** of the multiset.
- Given the multiset {*a*, *a*, *b*, *b*, *b*, *b*, *c*}, the
  - *multiplicities* of the **members** ***a***, ***b***, and ***c*** are **2**, ***4*** and **1** respectively
  - the *cardinality* of the **multiset** is 7.

Multiset Operations: Union

- We define operations on multisets that mirror the operations on sets.
- ***Union*** of multiset A and B:
  - Contains elements from A or B.
  - The number of occurrences for each element x equals to

    ***max* (f(x), g(x))**

  where f(x) and g(x) is number of occurrences of x in A and B respectively.
- Example:

  {*1*, **3**, 6, *1*, **3**, 1} U {1, **7**, ***6***, 3, ***6***} = {*1, 1, 1*, ***3, 3***, ***6, 6***, **7**}

Multiset Operations: Intersection

- **Intersection** of multiset A and B
  - ➢Contains elements that must be in BOTH A and B
  - ➢The number of occurrences for each element x is

$$\boldsymbol{\mathit{min}(f(x), g(x))}$$

where f(x) and g(x) are the number of occurrences of x in A and B respectively.

- Example:

{1, 3, 6, 1, 3, 1} ∩ {1, 7, 6, 1, 3, 6} = {1, 1, 3, 6}

Multiset Operations: Sum

- ***Sum*** of multiset A and B:
  - Contains elements that must be in BOTH A and B
  - The number of occurrences for each element x equals:

$$f(x) + g(x)$$

where f(x) and g(x) are number of occurrences of x in A and B respectively.

- Example:

{1, 3, 6, 1, 3, 1} + {1, 7, 6, 3, 6} = {1, 1, 1, 1, 3, 3, 3, 6, 6, 6, 7}

Other multiset operations can be similarly defined.

Computer Representation of Multisets

- **Representing a Multiset of integers:**

  int[] multiSetInt = new int [10];

  Index 0 to 9 represent integers,

  - Each element of the array can be set as the number of occurrences of that integer in this multiset
  - Examples:
    - If integer 6 exists in multiSetInt for 3 times, the element at location 6 should be set as 3

      {2, 6, 7, 2, 2, 6}

    can be represented as

      {0, 0, 3, 0, 0, 0, 2, 1, 0, 0}

Multiset implementation: Union

S1	0	1	2	3	4	5	6	7	8	9
	1	0	0	2	0	3	0	0	2	0

S2	0	1	2	3	4	5	6	7	8	9
	2	1	0	0	0	1	0	0	0	3

Let S3 be the multiset Union of S1 and S2:

S3[i] = max(S1[i], S2[i]);

Where i = 0, 1, …, 9;

S3	0	1	2	3	4	5	6	7	8	9
	2	1	0	2	0	3	0	0	2	3

Multiset implementation: Intersection

S1	0	1	2	3	4	5	6	7	8	9
	1	0	0	2	0	3	0	0	2	0

S2	0	1	2	3	4	5	6	7	8	9
	2	1	0	0	0	1	0	0	0	3

Let S3 be the multiset Intersection of S1 and S2:
S3[i] = min(S1[i], S2[i]);
Where i = 0, 1, …, 9;

S3	0	1	2	3	4	5	6	7	8	9
	1	0	0	0	0	1	0	0	0	0

Multiset implementation: Sum

S1	0	1	2	3	4	5	6	7	8	9
	1	0	0	2	0	3	0	0	2	0

S2	0	1	2	3	4	5	6	7	8	9
	2	1	0	0	0	1	0	0	0	3

Let S3 be the multiset Sum of S1 and S2:
S3[i] = S1[i] + S2[i];
Where i = 0, 1, …, 9;

S3	0	1	2	3	4	5	6	7	8	9
	3	1	0	2	0	4	0	0	2	3

Multiset implementation: Add an Element

■

S1	0	1	2	3	4	5	6	7	8	9
	2	0	4	1	0	0	0	2	0	0

Add 5 to S1:

S1	0	1	2	3	4	5	6	7	8	9
	2	0	4	1	0	**1**	0	2	0	0

Add 7 to S1:

S1	0	1	2	3	4	5	6	7	8	9
	2	0	4	1	0	1	0	**3**	0	0

Multiset implementation: Remove an Element

-

S1	0	1	2	3	4	5	6	7	8	9
	2	0	4	1	0	0	0	2	0	0

Remove 5 from S1: no change

S1	0	1	2	3	4	5	6	7	8	9
	2	0	4	1	0	**0**	0	2	0	0

Remove 7 from S1: change

S1	0	1	2	3	4	5	6	7	8	9
	2	0	4	1	0	1	0	**1**	0	0

Conclusion

- Sets and Multisets (finites)
  - Different functions to manage finite sets of data
    - Finite sets
  - From the formal definition to the use with computers
    - With arrays

Questions ?

- Reading
  - Csci 115 book – Section 1.2
  - Part VIII. B Sets, Introduction to Algorithm, 3rd Edition