

# Math Primer

Andrzej Wąsowski  
wasowski@itu.dk

February 12, 2007

Please, report all the bugs you can spot to the author!

**Acknowledgments:** Inspired by the appendix of Cormen et al. *Introduction to Algorithms*, and by some materials kindly provided by Volodya Shavrukov.

## Ingredients

A non exhaustive selection of mathematical concepts useful in an introductory algorithms course.

Ingredients: basic set theory, relations, equivalence relations, functions, important functions, graphs, trees, combinatorics, basic sums, probability

## 1 Basic Set Theory

$x \in S$  means  $x$  is a member of  $S$

$x \notin S$  means  $x$  is not a member of  $S$

Examples:

1.  $S = \{1, 2, 3\}$ ,  $2 \in S$ ,  $4 \notin S$

2.  $\emptyset$  — an empty set, for every  $a$  we have  $a \notin \emptyset$
3.  $\mathbb{Z}$  — the set of all integers,  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$
4.  $\mathbb{R}$  — all real numbers, cannot be counted

Sets are unordered:  $\{1, 2, 3\} = \{3, 2, 1\} = \{1, 2, 3, 1\}$

**Definition 1. Set inclusion:**  $A \subseteq B$  if every  $a \in A$  implies  $a \in B$

Example:  $\{1, 2\} \subseteq \{1, 2, 3\}$ .

**Definition 2. [-]** The set of all subsets of set  $S$  is called the powerset of  $S$ , denoted  $2^S$  or  $\mathcal{P}(S)$ .

[-] Example:  $2^{\{a,b\}} = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$

**Definition 3. A cartesian product** of two sets  $A, B$  denoted  $A \times B$  is the set of all ordered pairs, such that the first element of the pair is an element of  $A$  and the second is an element of  $B$ :

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$$

Example:  $\{a, b\} \times \{a, b, c\} = \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c)\}$

## 1.1 Set Operations

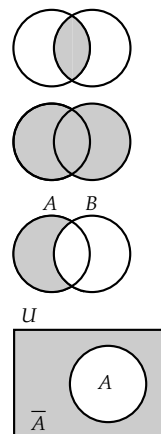
For any two sets  $A, B$  define

[intersection]  $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$

[union]  $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$

[difference]  $A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}$

[complement]  $\bar{A} = \{x \mid x \in U \text{ and } x \notin A\} = U \setminus A$



where  $U$  is the universe (set of all elements discussed) and  $A \subseteq U$ .

## 1.2 Set Laws

[-] Empty set:  $A \cap \emptyset = \emptyset$ ,  $A \cup \emptyset = A$

[-] Idempotence:  $A \cap A = A$ ,  $A \cup A = A$

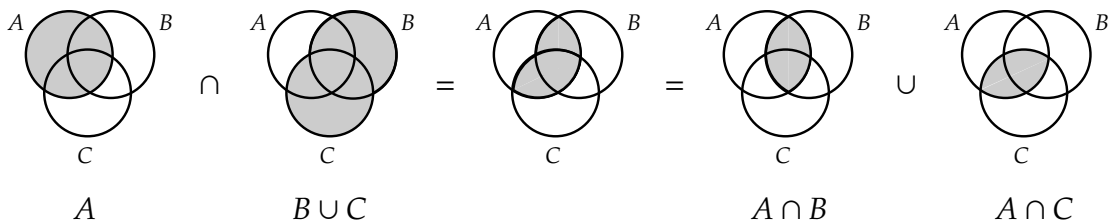
[-] Commutativity:  $A \cap B = B \cap A$ ,  $A \cup B = B \cup A$

Associativity:  $A \cap (B \cap C) = (A \cap B) \cap C$

Distributivity:  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Proof of the latter distributive law using Venn diagrams:



Venn diagrams only work up to three sets. Use definitions of operations to make proofs that cannot be made with Venn diagrams (or use definitions always).

[-] Absorption:  $A \cap (A \cup B) = A$ ,  $A \cup (A \cap B) = A$

[-] Complement Laws:  $\overline{\overline{A}} = A$ ,  $\overline{A} \cap A = \emptyset$ ,  $\overline{A} \cup A = U$

De Morgan's Laws:  $\overline{A \cap B} = \overline{A} \cup \overline{B}$   
 $\overline{A \cup B} = \overline{A} \cap \overline{B}$

**Exercise.** Prove de Morgan's laws using Venn diagrams.

## 1.3 Set Partition

**Definition 4.** Two sets  $A, B$  are **disjoint** if they have no elements in common:  $A \cap B = \emptyset$

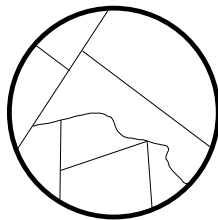
**Definition 5.** A collection  $\mathcal{S} = \{S_1, \dots, S_n\}$  forms a **partition of a set**  $S$  iff

1. All  $S_i$  are nonempty.
2. All  $S_i$  are pairwise disjoint:  $i \neq j \implies S_i \cap S_j = \emptyset$ .
3. Their union gives  $S$ :

$$S = S_1 \cup \dots \cup S_n = \bigcup_{S_i \in \mathcal{S}} S_i = \bigcup_{i=1}^n S_i$$

[−] Note: each element of  $S$  appears in exactly one  $S_i$ :

Partitioning into 8 classes:



## 1.4 Set Cardinality

$|S|$  — number of elements in set  $S$  (cardinality of  $S$ )

Properties:

$$|\emptyset| = 0$$

$$|A \cup B| = |A| + |B| - |A \cap B| \leq |A| + |B|$$

For  $A, B$  disjoint:  $|A \cup B| = |A| + |B|$

$$|2^S| = 2^{|S|}$$

$$|A \times B| = |A| \cdot |B|$$

## 2 Relations

**Definition 6.** An  $n$ -ary relation  $R$  on sets  $A_1, \dots, A_n$  is a subset of  $A_1 \times A_2 \times \dots \times A_n$ .

If  $n = 2$  then  $R$  is a binary relation.

$(a, b) \in R$  is typically written  $aRb$ .

Example: " $<$ " relation on natural numbers. We write  $1 < 2$  instead of  $(1, 2) \in <$ .

Binary relations over  $A \times A$

- $R \subseteq A \times A$  is **reflexive** iff for every  $a \in A$  it holds that  $aRa$   
Example: " $\leq$ " on natural numbers, but not " $<$ " on natural numbers.
- $R \subseteq A \times A$  is **irreflexive** iff for every  $a \in A$  it holds that  $(a, a) \notin R$   
Example: " $<$ " on natural numbers, but not " $\leq$ " on natural numbers.
- $R \subseteq A \times A$  is **symmetric** iff for every pair  $(a, b) \in A \times A$ ,  $aRb$  implies  $bRa$ .  
Example: " $=$ " on natural numbers, but not " $\leq$ ".
- $R \subseteq A \times A$  is **transitive** iff for every  $a, b, c \in A$  it holds that  $aRb$  and  $bRc$  together imply  $aRc$ .  
Example: " $=$ ", " $\leq$ ", " $<$ " on natural numbers
- [-]  $R \subseteq A \times A$  is **antisymmetric** iff for every pair  $(a, b) \in A \times A$  if  $aRb$  and  $bRa$  then also  $a = b$ .  
Example: " $\leq$ " on natural numbers

### 3 Equivalence Relations

**Definition 7.** A reflexive, symmetric and transitive relation is called an equivalence relation.

Example: “=” on natural numbers is an equivalence relation, but not “<” and “≤” (neither is symmetric and the former is not reflexive).

**Definition 8.** Let  $R$  be an *equivalence relation* on  $A$ . The equivalence class of element  $a \in A$  is a set  $[a] = \{ b \mid b \in A \text{ and } aRb \}$ .

Example:

$$R = \{(a, b) \mid a, b \in \mathcal{N} \text{ and } a + b \text{ is an even number.}\}$$

**Proposition 9.**  $R$  is an equivalence relation.

*Proof.*  $R$  is reflexive (as  $a + a$  is even for any natural number  $a$ ).

$R$  is symmetric (because  $a + b = b + a$ ).

$R$  is transitive: Assume that  $aRb$  and  $bRc$ . If  $a$  is even then  $b$  and  $c$  are also even, so  $a + c$  is even. If  $a$  is odd then  $b$  is odd, and also  $c$  must be odd, so  $a + c$  is even. □

Equivalence classes of  $R$ :  $[4] = \{0, 2, 4, 6, \dots\}$  and  $[3] = \{1, 3, 5, 7, \dots\}$ . There are no more distinct classes of  $R$ .

**Theorem 10.** 1. The equivalence classes of any equivalence relation  $R$  on  $A$  form a partition of  $A$ .

2. Any partition of  $A$  determines an equivalence relation on  $A$ , for which the sets in the partition are equivalence classes.

*Proof.* First show that equivalence classes form a partition:

1. Each equivalence class is non-empty as for every  $a \in [a]$  due to reflexivity of  $R$ .

2. They cover the entire set: each  $[a] \subseteq A$  and for each  $a \in A$  there exists a class containing  $a$  (indeed it is precisely  $[a]$ ). So  $\bigcup_{a \in A} [a] = A$ .
3. Equivalence classes are disjoint: if two different classes overlap then because of transitivity, they actually are the same class. Contradiction, so it cannot be that two different classes overlap.

The above three statements together mean that equivalence classes for a partition. Let us now turn to proving that a partition induces an equivalence class.

Let  $\mathcal{A} = \{A_1, \dots, A_n\}$  be a partition of  $A$ . We shall define an equivalence relation, whose equivalence classes are the  $A_i$  sets. Take:

$$R = \{ (a, b) \mid \text{there exists } i \text{ such that } a \in A_i \text{ and } b \in A_i \}$$

It is easy to show that  $R$  is reflexive, symmetric and transitive (exercise).

Equivalence classes of  $R$  are same as the  $A_i$  sets. First take any element  $a$  of any of the  $A_i$  sets. Observe that for any  $b \in A_i$  we have that  $b \in [a]$ , which means that  $A_i \subseteq [a]$ . At the same time for any  $b \in [a]$  we know that  $b \in A_i$  by definition of  $R$ , so  $[a] \subseteq A_i$ . By antisymmetry of set inclusion we conclude that  $A_i = [a]$ . □

## 4 Functions

**Definition 11.** A *function*  $f : A \rightarrow B$  is a binary relation on  $A$  and  $B$  such that for every  $a \in A$  there exists precisely one such  $b \in B$  that  $(a, b) \in f$ .

We typically write  $f(a) = b$  meaning  $(a, b) \in f$ . The set  $A$  is referred to as a **domain** of function  $f$ , while  $B$  is sometimes called a **codomain** of the function.

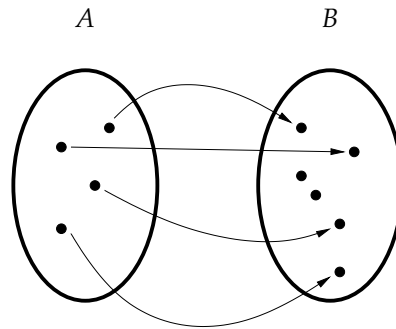


Figure 1: Function is a mapping.

Example:  $f : \mathbb{N} \rightarrow \{0, 1\}$ , where  $f = \{ (a, b) \mid a, b \in \mathbb{N} \text{ and } b = a \bmod 2 \}$

We would normally write:  $f(a) = a \bmod 2$

**Definition 12.** A function  $f : A \rightarrow B$  is a *surjection* if for every  $b \in B$  there exists an  $a \in A$  such that  $b = f(a)$ .

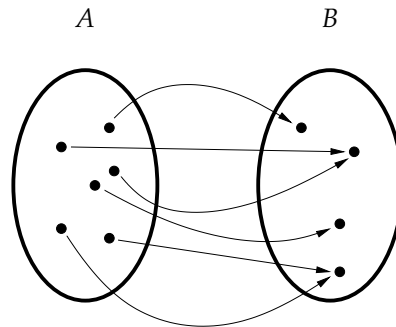


Figure 2: A surjection (all elements on the right hand side are pointed at).

**Q.** Which set has more elements on the above figure?  $A$  or  $B$  ?

Example:  $g : \mathbb{N} \rightarrow \mathbb{N}$ , where  $g(n) = n + 1$  is not a surjection (no value in the domain maps to zero in codomain).



**Definition 13.** A function  $f : A \rightarrow B$  is an **injection** if for distinct arguments it produces different values:  $a \neq a'$  implies  $f(a) \neq f(a')$ .

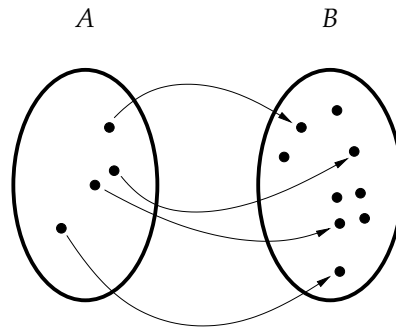


Figure 3: An injection (elements on the right hand side are pointed at by at most one arrow).

**Q.** Which set has more elements on the above figure?  $A$  or  $B$  ?

Example:  $g$  from earlier examples is an injection, while  $f$  is not.

**Definition 14.** A function  $f : A \rightarrow B$  is a **bijection** if it is both an injection and a surjection.

Injection defines a one-to-one correspondence between elements of domain and codomain.

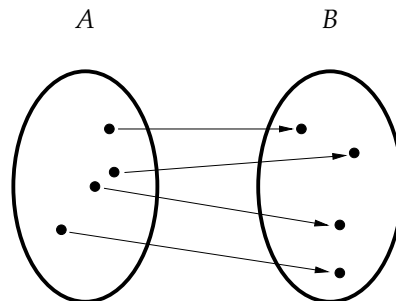


Figure 4: A bijection (all codomain elements are paired with domain elements).

**Q.** Which set has more elements on the above figure?  $A$  or  $B$  ?

Example: Neither  $f$  nor  $g$  are bijections. Function  $h : \mathbb{Z} \rightarrow \mathbb{Z}$ , where  $h(x) = -x$  is a bijection.

**Exercise.** Is the function  $f(x) = x + 1$  bijective when the domain and the codomain are  $\mathbb{N}$ ?

## 5 Common Functions

Floor  $\lfloor x \rfloor$  for a real number  $x$  is the greatest integer less than or equal to  $x$ .

Ceiling  $\lceil x \rceil$  for a real number  $x$  is the least integer greater than or equal to  $x$ .

### Exponentials

$$a^0 = 1 , \tag{1}$$

$$a^{-1} = \frac{1}{a} , \tag{2}$$

$$(a^m)^n = a^{mn} , \tag{3}$$

$$a^m a^n = a^{m+n} , \tag{4}$$

**Logarithms** A logarithm of  $a$  with base  $b$  ( $b \neq 1$ ), written  $\log_b a$ , is a number  $c$  such that  $b^c = a$ .

Assume that  $a, b, c > 0$ .

$$\log_c(ab) = \log_c a + \log_c b , \tag{5}$$

$$\log_b a^n = n \log_b a , \tag{6}$$

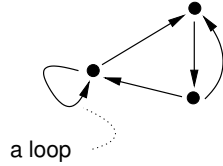
$$\log_b a = \frac{\log_c a}{\log_c b} \tag{7}$$

We typically mean a logarithm with base 2, when we drop the base writing  $\log a$  (sometimes also  $\lg a$ ). Some authors mean a logarithm with base 10, when writing  $\log a$ . Note that in algorithmic this causes no confusion, in most of the cases as  $\log_2 a$  is only by a constant factor different from  $\log_{10} a$ .

Similarly the natural logarithm of  $a$ , written  $\ln a$  (which is  $\log_e a$ ), differs only by a constant factor from other logarithms, too.

# 6 Graphs

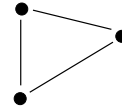
## 6.1 Directed Graphs



A directed graph (a *digraph*)

**Definition 15.** A directed graph  $G$  is a pair  $G = (\mathbf{V}, \mathbf{E})$ , where  $\mathbf{V}$  is a set of vertices, and  $\mathbf{E}$  is a binary relation on  $\mathbf{V}$  (edges)

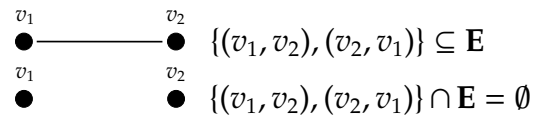
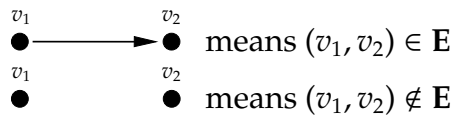
## 6.2 Undirected Graphs



An undirected graph

**Definition 16.** An undirected graph  $G$  is a pair  $G = (\mathbf{V}, \mathbf{E})$ , where  $\mathbf{V}$  is a set of vertices, and  $\mathbf{E}$  is a symmetric and irreflexive binary relation on  $\mathbf{V}$  (edges)

A formal interpretation of pictures:

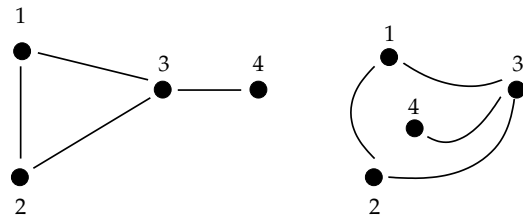


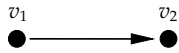
Our text book (RS) often uses  $E$  and  $V$  to denote number of edges and vertices:

$E = |\mathbf{E}|$  and  $V = |\mathbf{V}|$

$E = \frac{1}{2} |\mathbf{E}|$  and  $V = |\mathbf{V}|$

**Note:** Graph is not a picture! These pictures show the same undirected graph.



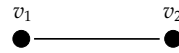


$(v_1, v_2)$  is incident *from*  $v_1$  (leaves  $v_1$ )

$(v_1, v_2)$  is incident *to*  $v_2$  (enters  $v_2$ )

out-deg  $v$  — *out-degree*, the number of edges leaving  $v$

in-deg  $v$  — *in-degree*, the number of edges entering  $v$



$\{v_1, v_2\}$  is incident *on* both  $v_1$  and  $v_2$

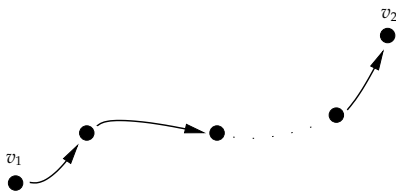
deg  $v$  — *degree*, the number of edges incident on  $v$

A property:

$$2E = |\mathbf{E}| = \sum_{v \in V} \text{deg } v \quad (8)$$

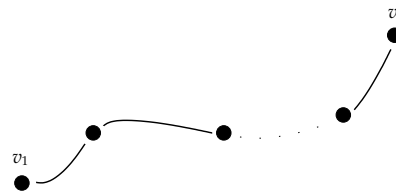
Q. Why?

### 6.3 Directed Paths




A path *from*  $v_1$  *to*  $v_2$

### 6.4 Undirected Paths



A path *between*  $v_1$  and  $v_2$

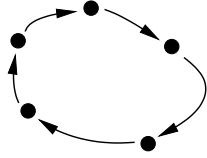
A graph  $G$  is *connected* if there is a path between any two of its vertices.

A path (both directed and undirected) is called a *simple path* if each of its vertices occurs in it at most once (so  cannot happen).

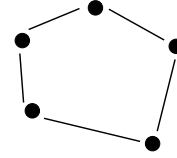
## 6.5 Directed Cycles

## 6.6 Undirected Cycles

A cycle is a path with both the beginning and the end in the same vertex.



A directed cycle



An undirected cycle

A directed graph not containing any cycles is called a *directed acyclic graph*

An undirected graph not containing any cycles is called a *forest*

If a forest is connected then we call it a *tree*

## 7 Trees

Trees look like this:

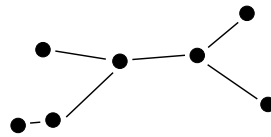
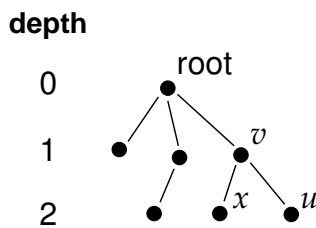


Figure 5: A tree

Trees are often *rooted*



ancestors of  $u$ :  $\{v, \text{root}\}$   
 $u$  is a *descendant* of root  
 $v$  is a *parent* of  $u$   
 $u$  is a *child* of  $v$   
 $x$  is a *sibling* of  $u$   
 $u$  is a *leaf* (so is  $x$ )  
 $v$  is an *internal node*

Figure 6: A rooted tree (the same as before, but with a designated root node)

**Definition 17.** A rooted tree such that each of its nodes has at most two children is called a binary tree. A binary tree is a complete binary tree of depth  $n$  if all its leaves have depth  $n$  and all its internal nodes have exactly two children.

We often denote a complete binary tree of depth  $n$  as  $T_n$ . Observe how wide the tree becomes with the growth of  $n$ .

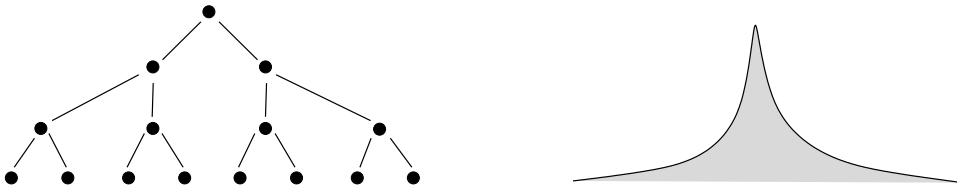


Figure 7:  $T_3$  on the left, a shape of  $T_n$  for large  $n$  on the right.

Q. How many leaves does  $T_n$  have?

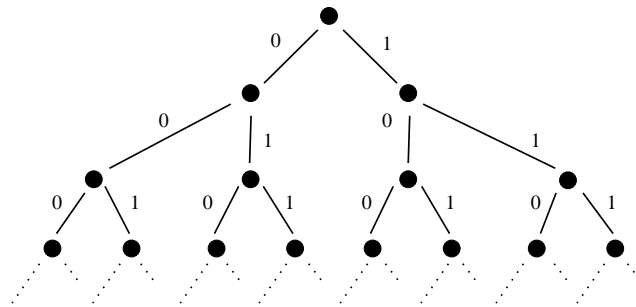


Figure 8: Any leaf is uniquely determined by a sequence of  $n$  bits.

So there is a bijection between the set of leaves of  $T_n$  and the set of all binary strings of length  $n$  (or in other words there is equally many of each).

There are  $2 \cdot 2 \cdot 2 \cdot \dots \cdot 2 = 2^n$  different binary sequences of length  $n$ .

## 8 Combinatorics & Basic Sums

Let  $X = \{a_1, \dots, a_n\}$

**Q.** How many subsets does  $X$  have?

Each subset is uniquely identified by a sequence of  $n$  bits, saying which elements are in and which are not. A bijection again:  $2^n$ .

**Q.** How many nodes in  $T_n$ ? (previously we have asked about leaves only)

$$2^n + 2^{n-1} + \dots + 2^1 + 2^0 = \frac{2^{n+1} - 1}{2 - 1} = 2^{n+1} - 1 \quad (9)$$

(computed using a formula for a sum of a geometric series)

**Geometric series** A sequence of numbers of the form  $x^0, x^1, x^2, x^3, \dots$  is called a *geometric series*. The sum of the first  $n$  elements can be computed using the following formula (for  $x \neq 1$ ):

$$x^n + x^{n-1} + \dots + x^1 + x^0 = \frac{x^{n+1} - 1}{x - 1} \quad (10)$$

Why? Take the sum and multiply it by  $(x - 1)$ . After simplifying a simple closed form is obtained, equal to the sum multiplied by  $(x - 1)$ .

$$\begin{aligned} (x^n + x^{n-1} + \dots + x^1 + x^0)(x - 1) &= \\ &= x^{n+1} + x^n + \dots + x^2 + x^1 - (x^n + x^{n-1} + \dots + x^1 + x^0) = \\ &= x^{n+1} - 1 \quad (11) \end{aligned}$$

Divide by  $(x - 1)$  and you are done.  $\square$

Incidentally if  $|x| < 1$  the value of the infinite sum is finite. In such case:

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad (12)$$

**Permutations** Let  $X = \{a_1, \dots, a_n\}$

A *permutation* of  $X$  is an ordered sequence of all elements of  $X$ .

Example. Let  $X = \{a, b, c\}$

Permutations of  $X$  are  $abc, bac, cab, acb, bca, cba$  (6 altogether).

**Definition 18.** Let  $k \leq n$ . A  $k$ -permutation of  $S = \{a_1, \dots, a_n\}$  is an ordered sequence of  $k$  distinct elements of  $S$ .

Example. Let  $X = \{a, b, c, d\}$ . 2-permutations of  $X$  are

$ab, ba, ca, da$

$ac, bc, cb, db$

$ad, bd, cd, dc$  (12 altogether)

We denote the number of  $k$ -permutations of an  $n$ -element set by  $P_k^n$ .

$$P_k^n = n(n-1)(n-2) \cdots (n-(k-1)) = \frac{n(n-1) \cdots 1}{(n-k)(n-k-1) \cdots 1} = \frac{n!}{(n-k)!} \quad (13)$$

$$S_n := P_n^n = n! \quad (14)$$

**Combinations.** A  $k$ -combination of an  $n$ -element set  $X$  is a  $k$ -element subset of  $X$ .  $C_k^n$  denotes the number of different  $k$ -combinations of an  $n$ -element set.

Example.  $X = \{a, b, c, d\}$ , 2-combinations of  $X$  are:  $\{a, b\}, \{b, c\}, \{a, c\}, \{b, d\}, \{a, d\}, \{c, d\}$  (6 altogether).



**Q.** How many?

Since

$$P_k^n = C_k^n S_k \quad (15)$$

so

$$C_k^n = \frac{P_k^n}{S_k} = \frac{n!}{(n-k)!k!} \quad (16)$$

**Arithmetic Series** The following sum is known as an arithmetic series:  $1 + 2 + 3 + 4 + \dots + \dots$ . The sum of the first  $n$  elements can be computed using the following formula:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} \quad (17)$$

It can be verified easily using mathematical induction.

## 9 Probability

An *elementary event* is a possible single outcome of an experiment. All elementary events are different and mutually exclusive (it cannot be that a single experiment results in two elementary events).

A set of all elementary events  $S$  for a given experiment is called a *sample space*.

An *event*  $A$  (note: not elementary) is a subset of the sample space, so  $A \subseteq S$ .

A function  $P : 2^S \rightarrow [0; 1]$  is a probability distribution iff

1.  $P(A) \geq 0$  for any event  $A \subseteq S$ .
2.  $P(S) = 1$
3.  $P(A \cup B) = P(A) + P(B)$  for any mutually exclusive events  $A \subseteq S, B \subseteq S$ .

For an event  $A$  the value of  $P(A)$  is called the *probability* of  $A$ .

Probability of an impossible event  $\emptyset$  is zero.

For two independent experiments  $S_1, S_2$  the probability of an event that the first one results in an event  $A_1 \subseteq S_1$  and the second results in  $A_2 \subseteq S_2$  equals  $P(A_1) \cdot P(A_2)$ .

A *random variable*  $X$  is a function from the sample space to the real numbers  $\mathbb{R}$ , assigning a real number with each possible outcome of an experiment.

For a random variable  $X$  and a number  $x$  we define  $X = x$  to be an event  $\{s \in S \mid X(s) = x\}$ . The probability of such an event is:

$$P\{X = x\} = \sum_{\{s \in S \mid X(s) = x\}} P(s) . \quad (18)$$

The *expected value*  $EX$  of a random variable  $X$  is defined by

$$EX = \sum_x x \cdot P\{X = x\} . \quad (19)$$