# DISCRETE STRUCTURES IN BIOMATHEMATICS

## An introduction for application

**Dietmar CIESLIK**
*Ernst-Moritz-Arndt University*
*Greifswald, Germany*

# CONTENTS

# PREFACE

> Philosophy is written in this grand book of the universe, which stands continually open to our gaze.... It is written in the language of mathematics.

Galileo Galilei

> One of the foundations of the mathematical method is that knowledge leads to more knowledge.

Michael Meyerson

Roughly speaking: Mathematics can be concerned as the essentially scientific part of any thoery. When investigating a ”real world problem” we make a lot of assumptions. The logical combination of these assumptions yields hints to the solution of the problem. Mathematics gives the possibility to order and to verify scientific facts.

Discrete Mathematics devoted to the study of discrete objects, these are

- a finite or countable set of distinct and unconnected elements; which are
- separated and discontiuous.

Discrete mathematics is used whenever objects are counted, when relationships between finite or countable sets are studied, and when processes involving a finite number of steps are analyzed.

Branches of discrete mathematics are

- **Combinatorics**
  which is the "Art of Counting".
  Many of the problems can be phrased in the form "How many ways...?", "Does there exist an object such that...?" or "Can we construct... ?"

- **Discrete probability and information theory**
  Probability is the branch of mathematics that deals with possible outcomes of events and their relative likelihoods.
  Information theory is concerned with discovering the laws of governing systems designed to communicate or manipulate information.

- **Graph Theory**
  which deals with binary relations.

- **Discrete Optimization**
  A general optimization problem is for a given configuration is to find an object, fulfilling some predetermined requirements and minimizes a given objective function.

More and more discrete structures are used in genetics, biochemistry, evolution, agriculture, experimental design and other parts of modern biology. Here, the results are very powerful and the research frontier are perhaps more accessible than in some more traditional areas of applied mathematics.
This book has been written for students reading biology, biochemistry or similar subjects. The aim in this under-graduate-level text is to outline the key mathematical concepts that underpin the important questions in discrete mathematics. In any case we will give examples for modelling structures and processes of biology with help of such objects.
This book originates from lectures and seminars given by the author at both Greifswald University and University of Science, Hanoi.

**Acknowledgements.**

I thank V.Liebscher and E.Rost, who gave me helpful advice on how to write this book. I thank my student D. Schlüter for proof reading the manuscript. Moreover, I thank my colleagues W. Girbardt and W. Passauer for helpful technical support.

The author accepts full responsibility for any mistakes that may have occured.

# 1

# BASICS

Set theory, founded by Cantor in the second half of the 19th century, has profoundly transformed mathematics. It is the foundation of modern mathematics.

## 1.1 SETS

A set is a collection of distinct objects. In contrast to a sequence of objects, a set is unordered. Usually, but not exclusive, we refer to the objects in a set as the elements. These elements may be sets themselves.
A set $S'$ is a subset of a set $S$, written $S' \subseteq S$, if every element of $S'$ is also an element of $S$. A proper subset is a subset with fewer elements than the whole set. Two sets $S$ and $S'$ are equal if they contain the same elements. In other words

$$S = S' \text{ if and only if } S \subseteq S' \text{ and } S' \subseteq S, \tag{1.1}$$

which gives a method for proving the equality of sets.
Two sets with no common elements are called disjoint.

There are two specific sets: the empty set $\emptyset$ containing no elements, and the universe $\mathcal{U}$ of all objects currently under consideration, or more philosophically, the part of the world under discussion[1].

We now introduce the following operations for sets.

[1] The universe of discourse.

- The union of sets:

$$S \cup S' = \{x \in \mathcal{U} : x \in S \text{ or } x \in S'\}. \tag{1.2}$$

- The intersection of sets:

$$S \cap S' = \{x \in \mathcal{U} : x \in S \text{ and } x \in S'\}. \tag{1.3}$$

- The set-difference of sets:

$$S \setminus S' = \{x \in \mathcal{U} : x \in S \text{ but not } x \in S'\}. \tag{1.4}$$

- The symmetric difference of sets:

$$\begin{aligned} S \triangle S' &= (S \cup S') \setminus (S \cap S') \qquad (1.5) \\ &= (S \setminus S') \cup (S' \setminus S). \qquad (1.6) \end{aligned}$$

- The complement of a set:

$$S^c = \{x \in \mathcal{U} : x \notin S\}. \tag{1.7}$$

A split of the set $S$ is a set $\{A, B\}$ of two nonempty subsets of $S$ such that

(a) $A \triangle B = S$, or equivalently

(b) $A \cup B = S$ and $A \cap B = \emptyset$.

The following table shows the effect of the symmetric difference for a split $\{A, B\}$ of $S$.

| | $\emptyset$ | $A$ | $B$ | $S$ |
|---|---|---|---|---|
| $\emptyset$ | $\emptyset$ | $A$ | $B$ | $S$ |
| $A$ | $A$ | $\emptyset$ | $S$ | $B$ |
| $B$ | $B$ | $S$ | $\emptyset$ | $A$ |
| $S$ | $S$ | $B$ | $A$ | $\emptyset$ |

**Observation 1.1.1** *Let $R, S, T \subseteq \mathcal{U}$. Then the following properties hold.*

*(a) The associative laws:*

$$\begin{aligned} (R \cup S) \cup T &= R \cup (S \cup T), \qquad (1.8) \\ (R \cap S) \cap T &= R \cap (S \cap T). \qquad (1.9) \end{aligned}$$

*(b) The commutative laws:*

$$\begin{aligned} R \cup S &= S \cup R, && (1.10) \\ R \cap S &= S \cap R. && (1.11) \end{aligned}$$

*(c) The distributive laws:*

$$\begin{aligned} R \cup (S \cap T) &= (R \cup S) \cap (R \cup T), && (1.12) \\ R \cap (S \cup T) &= (R \cap S) \cup (R \cap T). && (1.13) \end{aligned}$$

*(d) The empty set is the neutral element for the union:*

$$S \cup \emptyset = \emptyset \cup S = S. \qquad (1.14)$$

*(e) The universe is the neutral element for the intersection:*

$$S \cap \mathcal{U} = \mathcal{U} \cap S = S. \qquad (1.15)$$

*(f) The double-complement law:*

$$(S^c)^c = S. \qquad (1.16)$$

It can be easily verified that these properties are almost direct consequences of the definitions of the set operations. A more difficult property is given in the next theorem.

**Theorem 1.1.2** *(De Morgan's law) Let $R$ and $S$ be subsets of a universe $\mathcal{U}$. Then*

$$\begin{aligned} (R \cup S)^c &= R^c \cap S^c, && (1.17) \\ (R \cap S)^c &= R^c \cup S^c. && (1.18) \end{aligned}$$

*Proof.* Consider an element $x \in (R \cup S)^c$. Then $x \notin R \cup S$, and hence $x \notin R$ and $x \notin S$. This implies $x \in R^c$ and $x \in S^c$. Thus $x \in R^c \cap S^c$ which means

$$(R \cup S)^c \subseteq R^c \cap S^c.$$

Similarly it can be shown that

$$(R \cup S)^c \supseteq R^c \cap S^c,$$

whence equality holds.

The proof of the second formula can be carried out similarly.

□

In general, one cannot list the elements of a, in particular infinite, set. Nor it is practical to list the elements of a very large finite set. To determine a set of either kind we specify a property $P$ shared by all of its elements and not belonging to any element not in the set:

$$S = \{x \in \mathcal{U} : x \text{ satisfies } P\}. \tag{1.19}$$

The logical universe of discourse defines the set by all objects which posses an attribute.

## 1.2 THE NUMBER OF OBJECTS IN A SET

With $|S|$ we denote the number of elements of the set $S$. We can count the elements of $S$ by finding a bijection (a one-to-one correspondence) between $S$ and $\{1, \ldots, n\}$.[2]

**Observation 1.2.1** *Let $R$ and $S$ be sets in a universe $U$.*

*(a)* $|S^c| = |U| - |S|$.

*(b)* $|R \setminus S| = |R| - |R \cap S|$.

*(c)* $|R \cup S| = |R| + |S| - |R \cap S|$.

*(d)* $|R \triangle S| = |R \cup S| - |R \cap S| = |R| + |S| - 2|R \cap S|$.

Consider the power set of a set:

$$\mathcal{P}(S) = \{S' : S' \subseteq S\}. \tag{1.20}$$

Each subset $S'$ of a set $S = \{x_1, \ldots, x_n\}$ can be uniquely described by a 0/1-sequence $b_1, \ldots, b_n$ of length $n$:

$$b_i = \begin{cases} 1 & : \quad x_i \in S' \\ 0 & : \quad \text{otherwise} \end{cases}$$

Obviously, there are $2^n$ such sequences. Hence, the power set contains more elements than the set itself:

[2] Here we assume that $S$ is a finite set; later we will discuss the counting of infinite sets.

**Observation 1.2.2**

$$|\mathcal{P}(S)| = 2^{|S|}. \tag{1.21}$$

Moreover, our consideration gives us a way to enumerate methodically the subsets, beginning with the empty set $\emptyset$, and then adding each successive element of $S$ to a copy of each of all the previously listed subsets. Let us start with the following way of encoding subsets, illustrating it on the subsets of $S = \{a, b, c\}$. We look at the elements one by one, and write down a "1" if the element occurs in the subset and "0" if it does not. Thus each subset corresponds with a binary sequence of length 3. Moreover, such sequences remind us of the binary representation of integers.

| subset | binary seqence | integer |
|:---:|:---:|---:|
| $\emptyset$ | 000 | 0 |
| $\{c\}$ | 001 | 1 |
| $\{b\}$ | 010 | 2 |
| $\{b, c\}$ | 011 | 3 |
| $\{a\}$ | 100 | 4 |
| $\{a, c\}$ | 101 | 5 |
| $\{a, b\}$ | 110 | 6 |
| $S = \{a, b, c\}$ | 111 | 7 |

We see that the subsets correspond to the numbers 0,. . .,7. What happens if we consider subsets of a set with $n$ elements? We have binary sequences of length $n$ and use the numbers $0, 1, \ldots, 2^n - 1$.

## 1.3 ORDERED PAIRS, RELATIONS, CORRESPONDENCES

A list with two elements is normally called an ordered pair. If the first element of the pair is $a$ and the second element is $b$, we write $(a, b)$. The characteristic properties of pairs is that $(a, b) = (a', b')$ if and only if $a = a'$ and $b = b'$.
A list of $n$ objects is called an $n$-tupel, written $(a_1, a_2, \ldots, a_n)$. The cross-product of $n$ sets $S_1, \ldots, S_n$ is the set of all $n$-tupels

$$S_1 \times \cdots \times S_n = \{(a_1, \ldots, a_n) : a_i \in S_i, i = 1, \ldots, n\}. \tag{1.22}$$

Briefly we write

$$S^n = \underbrace{S \times \cdots \times S}_{n\text{–times}}. \tag{1.23}$$

It will be helpful to understand an $n$-tupel $(a_1, a_2, \ldots, a_n)$ equivalently as an consecutive written form of the objects: $a_1 a_2 \ldots a_n$, called a word, a (finite) sequence or a string.

### 1.3.1 Words

Molecular data comes in the form of

- DNA sequences, which are molecules containing information. This information is stored in the sequence of nucleotides from an alphabet of four letters; or
- Proteins, which are the operational molecules, composed of sequences of amino acids from an alphabet of 20 letters; or
- RNA sequences, which stand between both and are composed of nucleotides from an alphabet of four letters.

The relationship between DNA, RNA and protein as described by the Central Dogma of Molecular Biology can be summed up as follows:

- Integral form: DNA makes RNA makes protein.
- Differential form: Changed DNA can make changed protein.

An alphabet $A$ is a nonempty and finite set of distinguished letters (or symbols). If $A$ contains exactly one letter, all further discussed concepts and problems are senseless or trivial, respectively. Hence, we assume that $A$ contains at least two elements. If $A$ contains exactly two letters it is called a binary alphabet.
Important examples of alphabets are:

(a) $A = \{0, 1\}$ is an alphabet which plays a central role in coding theory. Moreover, we consider a word of 0's and 1's as a description of some individual, perhaps a genetic sequence in which each entry may take on one of two possible values.

(b) $A = \{a, c, g, t\}$ is the alphabet which codes the nucleotides of a DNA molecule, where $a$ stands for adenine, $c$ for cytosine, $g$ for guanine and $t$ for thymine. A similar alphabet, namely $A = \{a, c, g, u\}$ is used for the nucleotides of RNA, where $u$ codes for uracil.
Derived from this alphabet there is a binary alphabet $A' = \{r, y\}$ in which $r$ codes for a purine ($a$ or $g$), and $y$ codes for a pyrimidine ($c$ or $t$).

(c) The amino acids commonly found in proteins are coded by the alphabet $A = \{ala, arg, \ldots, val\}$, where the letters abbreviat the amino acids alanine, arginine,...,valine. In the usual genetic code $|A| = 20$ amino acids are coded, namely

| | One-letter code | Three-letter code | Name |
|---|---|---|---|
| 1 | A | ala | alanine |
| 2 | C | cys | cysteine |
| 3 | D | asp | aspartic acid |
| 4 | E | glu | glutamatic acid |
| 5 | F | phe | phenylalanine |
| 6 | G | gly | glycine |
| 7 | H | his | histidine |
| 8 | I | ile | isoleucine |
| 9 | K | lys | lysine |
| 10 | L | leu | leucine |
| 11 | M | met | methionine |
| 12 | N | asn | asparagine |
| 13 | P | pro | proline |
| 14 | Q | gln | glutamine |
| 15 | R | arg | arginine |
| 16 | S | ser | serine |
| 17 | T | thr | threonine |
| 18 | V | val | valine |
| 19 | W | trp | tryptophan |
| 20 | Y | tyr | tyrosine |

(d) The English language needs 26 letters: A,B,...,Y,Z, and a letter for the empty space. German needs several letters more: Ä, Ö, Ü, ß.

A word over an alphabet $A$ is a finite sequence of letters from $A$. The length $|w|$ of the word $w$ is the number of letters composing it. We additionally define an empty word $\lambda$ of length 0.
Note that the description of a word contains a left-to-right order of the letters. We will write $w = a_1 a_2 \ldots a_n$ for a word $w$ consisting of the letters $a_1$, $a_2$, $\ldots$ $a_n$ in this order. The letter $a_i$ in the word is called the $i$th position.

We say that two words $w = a_1a_2 \dots a_n$ and $w' = b_1b_2 \dots b_m$ over the same alphabet are equal, and we write $w = w'$, if $n = m$ and $a_i = b_i$ for all $i = 1, \dots, n$.

Let $w = a_1a_2 \dots a_n$ and $w' = b_1b_2 \dots b_m$ be two words over the same alphabet $A$. The concatenation of $w$ and $w'$, written $ww'$, is the word $a_1a_2 \dots a_nb_1b_2 \dots b_m$ over $A$. Hence, $|ww'| = |w| + |w'|$. Moreover, we will write $w^k = \underbrace{w \dots w}_{k-\text{times}}$ and $w^0 = \lambda$ for each word $w$.
For instance human insulin is composed by two words (chains):

**A:** gly ile val glu gln cys cys thr ser ile cys ser leu tyr glu leu glu asn tyr cys asn.

**B:** phe val asn gln his leu cys gly ser his leu val glu ala leu tyr leu val cys gly glu arg gly phe phe tyr thr pro lys thr.

$A^n$ is the set of all words over $A$ with length exactly $n$. Clearly, $A^0 = \{\lambda\}$ and $A^1 = A$. The set

$$A^\star = \bigcup_{n \geq 0} A^n \tag{1.24}$$

contains all words over the alphabet $A$; equipped with concatenation as a binary operation it satisfies the following properties:

(i) Closure: For all $v, w \in A^\star$, $vw \in A^\star$;

(ii) Association: For all $u, v, w \in A^\star$, $(uv)w = u(vw)$;

(iii) Identity: For the unity $\lambda$ it holds that for any $v \in A^\star$ it is $v\lambda = \lambda v = v$.

## 1.3.2 Relations

Let $\mathcal{U}$ be a universe. A subset of $\mathcal{U}^2$ is called a relation over $\mathcal{U}$.
Two specific kinds of relations will play important roles in our further considerations.

The elements of a universe $\mathcal{U}$ are said to satisfy a partial order $\preceq$ if

(i) $\preceq$ is reflexive: For all $x \in \mathcal{U}$ it holds that $x \preceq x$;

(ii) $\preceq$ is antisymmetric: If $x \preceq y$ and $y \preceq x$ then $x = y$;

(iii) $\preceq$ is transitive: For any three elements $x, y$ and $z$, if $x \preceq y$ and $y \preceq z$ then $x \preceq z$.

The pair $(\mathcal{U}, \preceq)$ is called a partially ordered set, or shortly a poset.
$\preceq$ is called a linear order if, additionally,

**(iv)** For any two elements $x$ and $y$ of $\mathcal{U}$, $x \preceq y$, $x = y$ or $y \preceq x$.

It is customary to use the symbol $\prec$ to denote $\preceq$ and $\neq$.

There are several important examples of ordered universes:

(a) Let $\mathcal{U}$ be a family of sets, then the relation $\subseteq$ of set inclusion is a partial, but of course not linear, order.

(b) Let $A$ be an alphabet. If there is an order $\leq$ of the letters in $A$, then the set $A^n$ is endowed with the following partial order $\leq_H$ of the words: For $w = a_1a_2 \ldots a_n$ and $w' = b_1b_2 \ldots b_n$ from $A^n$ we put

1. $w \leq_H w'$ if and only if $a_i \leq b_i$ for all $i = 1, \ldots, n$; and
2. $w <_H w'$ if and only if $w \leq_H w'$ and $w \neq w'$.

(c) Let $A$ be an alphabet. If there is an order $<$ of the letters in $A$, the set $A^\star$ is endowed with the following linear order $<_L$ of the words, which is called the lexicographic order: For two words $w = a_1a_2 \ldots a_n$ and $w' = b_1b_2 \ldots b_m$ we define $w <_L w'$ if

1. $n < m$ and $a_1 = b_1, \ldots, a_n = b_n$; or
2. $a_1 = b_1, \ldots, a_k = b_k$ for $k < n, m$ and $a_{k+1} < b_{k+1}$.

We write $w \leq_L w'$ if $w <_L w'$ or $w = w'$.

A relation in $\mathcal{U}$ is called an equivalence relation $\sim$ if

(i) $\sim$ is reflexive: For all $x \in \mathcal{U}$ it holds that $x \sim x$;

(ii) $\sim$ is symmetric: If $x \sim y$ then $y \sim x$;

(iii) $\sim$ is transitive: For any three elements $x, y$ and $z$, if $x \sim y$ and $y \sim z$ then $x \sim z$.

Let $\sim$ be an equivalence relation, then we define the equivalence classes by

$$C_x = \{y \in \mathcal{U} : x \sim y\}. \tag{1.25}$$

The collection

$$C(\sim) = \{C_x : x \in \mathcal{U}\} \tag{1.26}$$

of all equivalence classes creates a partition of $\mathcal{U}$, that means:

(i) No member of $C(\sim)$ is empty;

(ii) Any two different members of $C(\sim)$ are disjoint;

(iii) The union of all members of $C(\sim)$ is $\mathcal{U}$.

Conversely, a partition $C = \{C_i : i \in I\}$ creates an equivalence relation $\sim$ by $x \sim y$ if and only if there is an index $i$ such that $x, y \in C_i$.
Altogether, we find the following key result.

**Theorem 1.3.1** *If $C$ is a partition of $\mathcal{U}$, then there is one and only one equivalence relation whose equivalence classes are the members of $C$.*

### 1.3.3 Correspondences

We define a correspondence from a set $X$ to a set $Y$ to be a set of ordered pairs whose first entries are in $X$ and whose second entries are in $Y$. A function or mapping is a correspondence from $X$ to $Y$ that associates with each element $x \in X$ a unique element in $Y$. We frequently use a letter such as $f$ to stand for a function and write $f(x)$ for the element $f$ associates with $x$. We say $f$ maps $x$ to $f(x)$. The set

$$\mathrm{im}f = \{y \in Y : \text{ there is an element } x \in X : f(x) = y\} \tag{1.27}$$

is called the image of $f$.

A function from $X$ to $Y$ is called a one-to-one function or injection if it associates different elements in $Y$ with different elements in $X$. The function is called onto or a surjection if each element of $Y$ is associated with an element of $X$. A one-to-one function from $X$ onto $Y$ is called a bijection.
If $f$ is a bijection from $X$ onto $Y$, then the correspondence $f^{-1}$ defined by

$$f^{-1} = \{(y, x) : (x, y) \in f\}, \tag{1.28}$$

is a bijection from $Y$ onto $X$, which is called the inverse correspondence.

Let $f : X \to Y$ be a function from the set $X$ to a set $Y$, and let $g : Y' \to Z$ be a function from the set $Y' \supseteq Y$ to a set $Z$. Then the composition of $f$ and $g$ is the function $g \circ f : X \to Z$ given by

$$g \circ f(x) = g(f(x)) \tag{1.29}$$

for all $x \in X$.

**Observation 1.3.2** *Let $f$ be a bijection from $X$ onto $Y$, and let $g$ be the inverse function to $f$. Then $f \circ g$ is the identity on $X$ and $g \circ f$ is the identity on $Y$.*

## 1.3.4 Permutations

A permutation is a bijection from a finite nonempty set $S$ onto itself. In general we assume that $S = \{1, 2, 3, \ldots, n\}$, whenever $S$ contains $n$ elements. A convenient way to express a permutation $\pi$ is to write $\pi$ in an array form as

$$\pi = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n) \end{pmatrix}. \tag{1.30}$$

Let $\pi$ and $\kappa$ be permutations, then we define its product $\pi \cdot \kappa$ by the composition $\pi \circ \kappa$, that means $\pi \cdot \kappa(i) = \pi(\kappa(i))$. In this sense, we also defined the inverse of a permutation.

There is another notation commonly used to specify permutations, the so-called cycle notation. Each cycle is created by the following sequence: Start with the number $i$; then $\pi(i)$, $\pi^2(i) = \pi(\pi(i))$, $\pi^3(i)$, $\ldots$ until $\pi^k(i) = i$ for some $k$. Notice that for $k = 1$ the cycle contains only $i$. For instance:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 3 & 7 & 6 & 5 & 1 & 9 & 8 \end{pmatrix} \tag{1.31}$$

is $\pi = (1247)(3)(56)(89)$.
This cycle notation can be extended to any permutation by saying that the permutation can be written as product of disjoint cycles.
A transposition is a cycle of length 2. Notice that

$$(a_1 a_2 a_3 \ldots a_n) = (a_1 a_n)(a_1 a_{n-1}) \cdots (a_1 a_3)(a_1 a_2) \tag{1.32}$$

so that every permutation can be written as a product of transpositions.

## 1.4 THE MULTIPLICATION PRINCIPLE

We start with the following observation: If one thing can be done in $\alpha$ ways and a second thing can be done in $\beta$ ways independent of how the first thing is done, then the two things can be done in $\alpha \cdot \beta$ ways. More generally we have considering the cross-product of sets $S_i$:

$$S_1 \times \cdots \times S_n = \{(x_1, \ldots, x_n) : x_i \in S_i\}. \tag{1.33}$$

This notation is predictable, since the size of $S_1 \times \cdots \times S_n$ is the product of the sizes of the $S_i$:

**Observation 1.4.1** *(The multiplication principle)*

$$|S_1 \times \cdots \times S_n| = |S_1| \cdots |S_n|. \tag{1.34}$$

The *proof* is easy to understand: for $n = 2$ consider the rectangle of all pairs $(a_i, b_j)$, $i = 1, \ldots, |S_1|$, $j = 1, \ldots, |S_2|$. For the general case use induction on the number $n$.

□

A first application is to determine the size of the hypercube: $Q_n = \{0, 1\}^n$. The multiplication principle immediately gives

**Theorem 1.4.2**

$$|Q_n| = 2^n. \tag{1.35}$$

The number $2^n$ increases very rapidly as $n$ increases.[3]

As a second application consider the set $A^n$ of all words over $A$ with length exactly $n$. Clearly,

[3] This observation is expressed in ancient legends. The most famous is the following: A wise man who taught the king to play chess. As compensation, he asked the king to give him just one grain of rice on the first square of the chess board, two grains of the next square, and so on, doubling the number of grains for each succesive square. The result was that all the king's depots did not contain enough rice. (How much was necessary?)

**Theorem 1.4.3**

$$|A^n| = |A|^n. \tag{1.36}$$

Moreover, we can count the number of functions from a finite set $X$ into a finite set $Y$.

**Theorem 1.4.4** *Let $X$ and $Y$ be finite sets with $|X| = m$ and $|Y| = n$. Then the number of functions from $X$ to $Y$ is $n^m$.*

*Proof.* List the elements of $X$ in some order. Then each function can be represented by a $m$-tupel of elements from $Y$.

□

## 1.5 THE ADDITION PRINCIPLE

Suppose some event can occur in $\alpha$ ways and a second event can occur in $\beta$ ways, and suppose both events cannot occur simultaneously. Then both events can occur in $\alpha + \beta$ ways.

**Observation 1.5.1** *(The addition principle)*
*For a collection of pairwise disjoint sets the following holds:*

$$|S_1 \cup \ldots \cup S_n| = |S_1| + \ldots + |S_n|. \tag{1.37}$$

$A^{\leq n}$ denotes the set of all words of length at most $n$, which is the union of $n$ pairwise disjoint sets

$$A^{\leq n} = A^0 \cup A^1 \cup \ldots \cup A^n. \tag{1.38}$$

The addition principle and 1.4.3 lead to

$$\begin{aligned} |A^{\leq n}| &= |A^0 \cup A^1 \cup \ldots \cup A^n| \\ &= |A^0| + |A^1| + \ldots + |A^n| \\ &= |A|^0 + |A|^1 + \ldots + |A|^n \\ &= 1 + |A| + \ldots + |A|^n. \end{aligned}$$

If we multiply this equation $|A|$-times with itself and then subtract the equation from the product we get

$$(|A|-1)\cdot|A^{\leq n}| = |A|^{n+1}-1. \tag{1.39}$$

Hence,

**Theorem 1.5.2** *Let $A$ be an alphabet of at least two letters. Then*

$$|A^{\leq n}| = \frac{|A|^{n+1}-1}{|A|-1} < |A|^{n+1}. \tag{1.40}$$

A very interesting example in biology is the following: The number of polypeptide sequences with at most 100 amino acids which can theoretically exist is

$$|\{\text{ ala },\ldots,\text{ val }\}^{\leq 100}|-1 = \frac{20^{101}-1}{19}-1 \approx 10^{130}, \tag{1.41}$$

which is a number greater than the number of atoms in the universe. Consequently only very few sequences describe proteins of present or ancient living entities.

Shortly, we discuss the important case that the sets are not disjoint:

$$|S_1\cup S_2| = |S_1|+|S_2|-|S_1\cap S_2|. \tag{1.42}$$

That is we "include" $S_1$ and $S_2$, and we "exclude" $S_1\cap S_2$. This follows from the fact that, when we add $|S_1|$ and $|S_2|$ elements, we have counted the elements from $S_1\cap S_2$ twice.
Similar for three sets: In the sum $|S_1|+|S_2|+|S_3|$ an element of $S_1\cap S_2$ is included at least twice; an element of $S_1\cap S_2\cap S_3$ three times. Hence,

$$\begin{aligned}|S_1\cup S_2\cup S_3| &= |S_1|+|S_2|+|S_3|\\ &\quad -|S_1\cap S_2|-|S_1\cap S_3|-|S_2\cap S_3|\\ &\quad +|S_1\cap S_2\cap S_3|. \end{aligned}\tag{1.43}$$

Such a formula is called an inclusion-exclusion formula, and we can generalize it to:

**Theorem 1.5.3** *Let $S_1,\ldots,S_n$ be a collection of sets. Then*

$$\begin{aligned}|\bigcup_{i=1}^{n} S_i| &= \sum_{i=1}^{n}|S_i| - \sum_{1\leq i<j\leq n}|S_i\cap S_j|\\ &\quad + \sum_{1\leq i<j<k\leq n}|S_i\cap S_j\cap S_k| \mp \ldots - (-1)^n|\bigcap_{i=1}^{n} S_i|.\end{aligned}\tag{1.44}$$

## 1.6 FACTORIALS

### 1.6.1 Placing objects in a row

How many ways are there of placing the three objects $x$,$y$ and $z$ in a row? There are six ways, namely $xyz$,$xzy$,$yxz$,$yzx$,$zxy$ and $zyx$.
Such linear arrangements of distinct objects are permutations, which number can easily be determined.

**Theorem 1.6.1** *There are*

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 \tag{1.45}$$

*ways to place $n$ objects in a row.*

*Proof.* There are $n$ possibilities for the first place, then $n-1$ for the second place, and so on until there is just one for the last place. So we get the assertion from the multiplication principle.

□

It is convenient to define $0! = 1$. Obviously, the function ! satisfies the following recurrence relation:

$$\begin{aligned} n! &= n \cdot (n-1)! \quad \text{for } n \geq 1; & (1.46) \\ 0! &= 1. & (1.47) \end{aligned}$$

### 1.6.2 Anagrams

The number of arrangements of the four letters in BALL is not $24 = 4!$, since we do not have four distinct letters to arrange. The letter L occurs twice, and we have to count $4!/2 = 12$. Generalizing this idea we solved a new type of problem by relating it to the previous enumeration principles:

**Observation 1.6.2** *If there are $n$ objects of $r$ types with $n_i$ of the $i$th type, $i = 1, \ldots, r$, where $n_1 + \ldots + n_r = n$, then there are*

$$\frac{n!}{n_1! \cdot n_2! \cdots n_r!} \tag{1.48}$$

*(linear) arrangements of the objects.*

RNA is a messenger molecule whose links are defined by DNA. The possible bases (letters) are adenine (a), cytosine (c), guanine (g) and uracil (u). A sequence (word) of bases encodes certain genetic information. It is an elementary problem of combinatorics to find the number of possible RNA sequences with certain link makeup.
Let $n_k$ with $k \in \{a, c, g, u\}$ be the number of bases of this type. Then $n = n_a + n_c + n_g + n_u$ is the length of the RNA, and in view of 1.6.2 we have $n!/n_a! \cdot n_c! \cdot n_g! \cdot n_u!$ of these sequences.

Genralizing this example and combining 1.4.3 and 1.6.2 we find

**Observation 1.6.3**

$$\sum_{n_1+n_2+\ldots+n_r=n} \frac{n!}{n_1! \cdot n_2! \cdots n_r!} = r^n. \tag{1.49}$$

## 1.7 METRIC SPACES

Distance is the mathematical description of the idea of proximity, and consequently, will play an important role in mathematics.

A pair $(X, \rho)$ is called a metric space if $X$ is a nonempty set of elements called the points, and $\rho : X \times X \to I\!R$ is a real-valued function satisfying:

(i) $\rho(x, y) \geq 0$ for all $x, y$ in $X$;

(ii) $\rho(x, y) = 0$ if and only if $x = y$;

(iii) $\rho(x, y) = \rho(y, x)$ for all $x, y$ in $X$; and

(iv) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for all $x, y, z$ in $X$ (triangle inequality).

Usually, such a function $\rho$ is called a metric.
We will say that the quantity $\rho(x, y)$ is the distance between the points $x$ and $y$.
The following variants of "metric approaches" will be also of interest:

- If $\rho$ satisfies (ii) only in the weaker form

  (ii') $\rho(x,x) = 0$ for all $x$ in $X$;

  we say that $\rho$ is a pseudometric.

- If the function $\rho$ satisfies the conditons (i),(ii') and (iii) it is called a dissimilarity.[4]

We will find distances for many sets and of great importance. Firstly, we consider the following examples.

(a) The Euclidean plane is defined in the affine plane with the Euclidean metric $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ between the points $(x_1, y_1)$ and $(x_2, y_2)$ derived from a norm $||.||$:

$$||(x,y)|| = \sqrt{x^2 + y^2}. \tag{1.50}$$

(b) Space with rectilinear distance: In the $d$-dimensional affine space the distance between $(x_1, \ldots, x_d)$ and $(y_1, \ldots, y_d)$ is defined as

$$\sum_{i=1}^{d} |x_i - y_i|. \tag{1.51}$$

(c) Using the binary operation $\triangle$ we find a metric for sets

**Observation 1.7.1** *$|S_1 \triangle S_2|$ is a metric.*

It is sufficient to show the triangle inequality.

$$S_1 \triangle S_2 \subseteq S_1 \triangle S_3 \cup S_3 \triangle S_2. \tag{1.52}$$

Moreover,

$$S_1 \triangle S_3 \cap S_3 \triangle S_2 = ((S_1 \cap S_2) \setminus S_3) \cup (S_3 \setminus (S_1 \cup S_2)). \tag{1.53}$$

That means, if an element is in $S_1 \triangle S_3 \cap S_3 \triangle S_2$, then it cannot be in $S_1 \triangle S_2$.

□

[4] We will give the reason for this name later. There are various measures of dissimilarity, and not all of them yield a metric, but many do.

(d) The following metric can be created over any set $X$ of points:

$$\rho(x,y) = \begin{cases} 0 & : \quad x = y \\ 1 & : \quad \text{otherwise} \end{cases}$$

(e) To consider the problem of reconstruction of evolutionary (phylogenetic) trees, we introduce so-called sequence spaces. These are metric spaces whose points are arbitrary words generated by letters from some (finite) alphabet, and the metric measuring "sameness" of the words: Let $v = a_1 a_2 \dots a_n$ and $w = b_1 b_2 \dots b_n$ be DNA-sequences of length $n$. We define the Hamming distance by the number of positions in which $v$ and $w$ disagree:

$$\rho_H((a_1, \dots, a_n), (b_1, \dots, b_n)) = |\{i : a_i \neq b_i \text{ for } i = 1, \dots, n\}|. \tag{1.54}$$

Let $(X, \rho)$ be a metric space. A set $S$ is called bounded, if there exists a real number $d$ such that

$$\rho(x,y) \leq d \tag{1.55}$$

for each pair $x, y \in S$. For a bounded set $S$ we define the diameter as

$$D(S) = \sup\{\rho(x,y) : x, y \in S\}. \tag{1.56}$$

## 1.8 INFINITE SETS

The concept of infinity has always fascinated philosophers and theologians, but that was avoided or met with open hostility throughout most of the history of mathematics. Only within the last two centuries mathematicians dealt with it head on and accepted infinity as a number.

How can we count the elements of an infinite set? We have to compare the sets; that means we ask for a bijective mapping between these sets.

### 1.8.1 Can a part be equal to the whole?

One dogma that we have to brush aside is the statement "A part is less than the whole". This is indisputately true for finite sets, but it loses its force when we try to apply it to infinite sets. Consider the following mapping:

$$f : I\!\!N \rightarrow I\!\!N : n \mapsto 2n. \tag{1.57}$$

This sets up a one-to-one correspondence between the set of natural numbers and a proper part of this set: the set of even numbers, which was already observed by Galilei.[5] Dedekind introduced the concept of infinite sets by the following definition.

> An infinite set is as one that can be placed into a one-to-one correspondence with a proper subset of itself.

### 1.8.2 Transfinite numbers

Cantor developed the idea of levels of infinity. To carry a notion of equal size of two finite or infinite sets $X$ and $Y$ we define that this is given if a bijective mapping from $X$ onto $Y$ exists. In other terms, the elements of $X$ and $Y$ may be paired with each other in such a way that to each element of $X$ there corresponds one and only one element of $Y$ and vice versa. This notation for finite sets coincides with the ordinary notation of equality of numbers.

**Observation 1.8.1** *The notion of equal size is an equivalence relation.*

*Proof.* We check the three properties of an equivalence relation.

(i) The identity is a bijective mapping.

(ii) For symmetry compare 1.3.2.

(iii) Let $f : X \to Y$ be a bijective mapping from $X$ to $Y$, and let $g : Y \to Z$ be a bijective mapping from $Y$ to $Z$. Then the composition of $f \circ g$ is also bijective.

□

Consequently, we can associate a number, called cardinal number, with every class of equal-sized sets. The cardinal numbers of infinite sets are called transfinite numbers.

[5] Here is the story of Hilbert's hotel: It is a hotel with an infinite number of rooms. All the rooms are full, but more guests are waiting outside. We amke space by the following operation: the guest occypying room 1 moves to room 2, the occypant from room 2 moves to room 4, and so on, all the way down the line, an infinite number of newcomer can be placed in the empty rooms.

### 1.8.3 Countable sets

We call sets with as many elements as the set of natural numbers countable sets. In other words, a set is called countable if it is infinite and its elements can be counted with the aid of the natural numbers. For example the set $\mathbf{\Gamma}$ of integers is countable:

| $I\!N$ | 0 | 1 | 2 | 3 | 4 | $\cdots$ | $2n-1$ | $2n$ | $\cdots$ |
|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{\Gamma}$ | 0 | 1 | -1 | 2 | -2 | $\cdots$ | $n$ | $-n$ | $\cdots$ |

That means the following function $f : I\!N \to \mathbf{\Gamma}$ is one-to-one and onto:

$$f(n) = \begin{cases} -\frac{n}{2} & : \ n \text{ even} \\ \frac{n+1}{2} & : \ \text{otherwise} \end{cases}$$

It is more difficult to show that the rational numbers are also countable. Obviously this is paradoxical: Between any two rational numbers we can still find infinitely many rational numbers. So it is quite unclear how we should go about numbering them.
First we prove that $I\!N^2$ is countable. Consider the following tabulation, which is called Cantor's first diagonal principle.

| $x \setminus y$ | 0 | 1 | 2 | 3 | 4 | $\ldots$ |
|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 3 | 6 | 10 | $\ldots$ |
| 1 | 2 | 4 | 7 | 11 | 16 | $\ldots$ |
| 2 | 5 | 8 | 12 | 17 | 23 | $\ldots$ |
| 3 | 9 | 13 | 18 | 24 | 31 | $\ldots$ |
| 4 | 14 | 19 | 25 | 32 | 40 | $\ldots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | |

That means at first we count all pairs $(x, y)$ with $x + y = 0$, then all pairs with $x + y = 1$, then with $x + y = 2$, and so on. The pair $(x, y)$ lies in position number $x$ between $(0, x + y)$ and $(x + y, 0)$. Before $(0, x + y)$ we have exactly

$$1 + 2 + \ldots + (x + y) = \frac{(x + y)(x + y + 1)}{2}$$

pairs. Hence,

**Theorem 1.8.2** *The function*

$$c : (x, y) \mapsto \frac{(x+y)(x+y+1)}{2} + x \tag{1.58}$$

*is a bijective mapping from* $I\!N^2$ *onto* $I\!N$.

The functions

$$\begin{aligned} l(n) &= n - \frac{1}{2}\left\lfloor \frac{\lfloor\sqrt{8n+1}\rfloor + 1}{2} \right\rfloor \left\lfloor \frac{\lfloor\sqrt{8n+1}\rfloor - 1}{2} \right\rfloor & (1.59) \\ r(n) &= \left\lfloor \frac{\lfloor\sqrt{8n+1}\rfloor + 1}{2} \right\rfloor - l(n) - 1 & (1.60) \end{aligned}$$

are the inverse mappings of $c$, which means

$$c(l(n), r(n)) = n. \tag{1.61}$$

We will omit the proof.
With 1.8.2 in mind, we have several considerations.

**Corollary 1.8.3** *For any integer* $n \geq 2$ *there exist a bijective mapping* $c^{(n)}$ *from* $I\!N^n$ *onto* $I\!N$.

*Proof.* Let $c$ be a bijective mapping from $I\!N^2$ onto $I\!N$, compare 1.8.2. We create $c^{(n)}$ by the following recursive equations:

$$\begin{aligned} c^{(2)} &= c, & (1.62) \\ c^{(n)}(x_1, \ldots, x_n) &= c^{(n-1)}(c(x_1, x_2), x_3, \ldots, x_n), & (1.63) \end{aligned}$$

$n = 3, 4, \ldots$.

□

For further applications we start with the observation that a countable set is the smallest of the infinite sets:

**Lemma 1.8.4** *Consider infinite sets.*

*(a) Any infinite set contains a countable set.*

*(b) An infinite subset of a countable set is also countable.*

*Proof.* (a): We can select a countable subset from an infinite set $S$ in the following way: Take any element $x_0$ from $S$. Clearly, we have not exhausted the elements of $S$ with the selection of $x_0$, so we can proceed to select a second element $x_1$. After that we select a third element $x_2$ and so on. We have thus extracted from $S$ a countable subset of indexed element.
(b) is an immediately consequence of (a).

□

**Corollary 1.8.5** *The set of all rational numbers is countable.*

*Proof.* Consider the correspondence

$$(x, y) \in I\!N^2 \mapsto \left\{ \begin{array}{ccl} \frac{x}{y} & : & y \neq 0 \\ x & : & \text{otherwise} \end{array} \right.$$

Then apply 1.8.4(b) and our introductionary example.

□

**Theorem 1.8.6** *Let $S_1, S_2, \ldots$ be a countable number of finite sets, then the union $S = \bigcup_i S_i$ is finite or countable.*

*Proof.* We define sets $R_1, R_2, \ldots$ where $R_i$ contains the elements of $S_i$ which do not belong to preceding sets, that means

$$\begin{aligned} R_1 &= S_1 && (1.64) \\ R_i &= S_i \setminus (S_1 \cup S_2 \cup \ldots \cup S_{i-1}) && (1.65) \end{aligned}$$

for $i \geq 2$. Then the $R_i$ are disjoint and $\bigcup_i R_i = S$.
Let

$$R_i = \{b_{i1}, b_{i2}, \ldots, b_{im_i}\}. \qquad (1.66)$$

If $S = \{b_{ij}\}$ is infinite, then we define a bijective function $f$ from $S$ onto the natural numbers by

$$f(b_{ij}) = m_1 + m_2 + \ldots + m_{i-1} + j. \qquad (1.67)$$

□

**Theorem 1.8.7** *A countable union of countable sets is countable.*

*Proof.* Let $S_1, S_2, \ldots$ be a countable number of countable sets, and suppose that $a_{i1}, a_{i2}, \ldots$ are the elements of $S_i$. We define sets $R_2, R_3, R_4, \ldots$ as follows:

$$R_k = \{a_{ij} : i + j = k\}. \tag{1.68}$$

Observe that each $R_k$ is a finite set and

$$\bigcup_k R_k = \bigcup_i S_i. \tag{1.69}$$

Then we apply 1.8.6.

□

### 1.8.4 The number of words

Remember our definition of words over an alphabet $A$. $A^n$ is the set of all words over $A$ with length exactly $n$. We know $|A^n| = |A|^n$, that means, that each of the sets $A^n$ is finite.
The set

$$A^\star = \bigcup_{n \geq 0} A^n \tag{1.70}$$

contains all words over the alphabet $A$. In any case, this is an infinite, but countable, set.[6] To see the countableness, we use 1.8.3 and 1.8.7. Additionally, we will give a direct method to count the words. First count the word $\lambda$, then the members of $A$ itself, then the words of length 2, and so on. More precisely, let $A = \{a_1, \ldots, a_n\}$, then

[6] Also for a one-element alphabet $A$ the set $A^\star$ is infinite: $A = \{|\}$, then $A^\star = \{\lambda, |, ||, |||, |^4, |^5, \ldots\}$.

| $I\!N$ | $A^\star$ |
|---|---|
| $0$ | $\lambda$ |
| $1$ | $a_1$ |
| $2$ | $a_2$ |
| $\vdots$ | $\vdots$ |
| $n$ | $a_n$ |
| $n+1$ | $a_1a_1$ |
| $n+2$ | $a_1a_2$ |
| $\vdots$ | $\vdots$ |
| $2n$ | $a_1a_n$ |
| $2n+1$ | $a_2a_1$ |
| $2n+2$ | $a_2a_2$ |
| $\vdots$ | $\vdots$ |
| $3n$ | $a_2a_n$ |
| $3n+1$ | $a_3a_1$ |
| $\vdots$ | $\vdots$ |
| $n^2+1$ | $a_na_1$ |
| $\vdots$ | $\vdots$ |
| $n^2+n$ | $a_na_n$ |
| $n^2+n+1$ | $a_1a_1a_1$ |
| $\vdots$ | $\vdots$ |

### 1.8.5 Uncountable sets

All the sets we have constructed so far have been countable. This naturally leads us to ask whether all infinite sets are countable. But the situation turns out to be more complicated than that; uncountable sets exist, and of more than one cardinality.

First we show, using Cantor's second diagonal principle

**Theorem 1.8.8** *The set of all (infinite) binary sequences is not countable.*

*Proof.* Assume that there is a counting of $\{0,1\}^\infty$ given by the following double infinite array:

| $\mathbb{N}$ | $\{0,1\}^\infty$ |
|---|---|
| 0 | $b_{00}, b_{01}, b_{02}, b_{03}, \ldots$ |
| 1 | $b_{10}, b_{11}, b_{12}, b_{13}, \ldots$ |
| 2 | $b_{20}, b_{21}, b_{22}, b_{23}, \ldots$ |
| 3 | $b_{30}, b_{31}, b_{32}, b_{33}, \ldots$ |
| $\vdots$ | $\vdots$ |

The sequence $b_0, b_1, b_2, \ldots$ with $b_i = 1 - b_{ii}$ cannot be in this table.

□

With 1.8.8 in mind, we have the following considerations.

**Observation 1.8.9** *The set of all real numbers is uncountable.*

Any set that can be brought into a bijective correspondence with the set of real numbers is called a continuum.[7]

### 1.8.6 Functions and the power set

$\mathcal{F}(X, Y)$ denotes the collection of all functions $f : X \to Y$.

**Theorem 1.8.10** *The set $\mathcal{F}(X, Y)$ contains more elements than $X$ whenever $Y$ contains at least two elements.*[8]

[7] Cantor's next discovery was a shock even to Cantor himself:
The real plane $\mathbb{R}^2$ has the same size as $\mathbb{R}$.
Consequently, all sets $\mathbb{R}^n$, $n \geq 1$ are of the same size.

For the proof, it will be sufficient to map all pairs $(x, y)$, $0 \leq x, y < 1$ bijectively onto the interval $[0, 1)$. Consider $(x, y)$ and write $x$ and $y$ in their digits:

$$\begin{aligned} x &= 0.a_1a_2a_3\ldots \\ y &= 0.b_1b_2b_3\ldots. \end{aligned}$$

We now create a number $z$ by ”mixing” the digits:

$$z = 0.a_1b_1a_2b_2a_3\ldots.$$

(There is a little problem in this correspondence. Do you see which?)
This result is not what we would expect from our idea of dimension. Dimension is not generally preserved by bijective mappings.

[8] The proof will show that this assertion is true for finite and infinite sets $X$.

*Proof.* First we show that there are as many functions in $\mathcal{F}(X,Y)$ as elements in $X$. Consider for each $x_0 \in X$ the function $f[x_0]$ defined by

$$f[x_0](x) = \begin{cases} y_1 & : \quad x = x_0 \\ y_2 & : \quad \text{otherwise} \end{cases}$$

where $y_1$ and $y_2$ are distinct elements of $Y$.
If $x_0 \neq x_1$ then $f[x_0] \neq f[x_1]$.
Now assume that there is a bijective mapping

$$x \in X \mapsto f[x] \in \mathcal{F}(X,Y). \tag{1.71}$$

Choose $y_x \in Y$ such that $y_x \neq f[x](x)$. The function $f$ defined by

$$f : x \mapsto y_x, \tag{1.72}$$

cannot be one of the function $f[x]$.

□

As exercise prove the following fact.

**Theorem 1.8.11** *The power set $\mathcal{P}(X) = \{X' : X' \subseteq X\}$ contains more elements than $X$.*

In view of these facts we have the following two important consequences:

- A largest cardinal number, both finite and transfinite, does not exist, since in both cases we start with a finite and infinite set $X$, respectively, and build the following infinite sequence of sets:

$$X, \mathcal{P}(X), \mathcal{P}(\mathcal{P}(X)), \mathcal{P}(\mathcal{P}(\mathcal{P}(X))), \ldots. \tag{1.73}$$

- Each language is a countable set.[9] In view of 1.8.7 a union of all "real-world" languages is countable. But, paying attention 1.8.11, the collection of all languages is uncountable, and hence not each language can be numbered.

[9] In the strict sense all practical languages are finite, this comes from the finiteness of the real world; compare [35] and [60].

## 1.9 ORDERS OF GROWING AND ASYMPTOTICS

In the following discussion we will use the phrase "on the order of" to express lower and upper bounds. For this purpose we introduce specific notations, called Landau symbols.

### 1.9.1 Landau symbols

Let $f$ and $g$ be functions from the positive integers into the real numbers. Then:

(i) The function $g(n)$ is said to be of order at least $f(n)$, denoted $\Omega(f(n))$, if there are positive constants $c$ and $n_0$ such that $g(n) \geq c \cdot f(n)$ for all $n \geq n_0$.

(ii) The function $g(n)$ is said to be of order at most $f(n)$, denoted $O(f(n))$, if there are positive constants $c$ and $n_0$ such that $g(n) \leq c \cdot f(n)$ for all $n \geq n_0$.

(iii) The function $g(n)$ is said to be of order $f(n)$, denoted $\Theta(f(n))$, if $g(n) = \Omega(f(n))$ and $g(n) = O(f(n))$. That is, $f(n)$ and $g(n)$ both grow at the same rate; only the multiplicative constants may be different.

This notation allows us to concentrate on the dominating term in an expression describing a lower or upper bound and to ignore any multiplicative constants. $O(f(n))$ is to be read "big-O of $f$ of $n$". Note that it is not an equation in the usual sense. It has to be read only from left to right.

**Theorem 1.9.1** *If $f(n)$ is a polynomial of degree $k$ that means*

$$f(n) = a_k x^k + a_{k-1} x^{k-1} + \ldots + a_1 x + a_0. \tag{1.74}$$

*Then*

$$f(n) = O(n^k). \tag{1.75}$$

*Proof.*

$$\begin{aligned} |f(n)| &= |a_k x^k + a_{k-1} x^{k-1} + \ldots + a_1 x + a_0| \\ &\leq |a_k x^k| + |a_{k-1} x^{k-1}| + \ldots + |a_1 x| + |a_0| \end{aligned}$$

$$
\begin{aligned}
&= |a_k|x^k + |a_{k-1}|x^{k-1} + \ldots + |a_1|x + |a_0| \\
&\leq (|a_k| + |a_{k-1}| + \ldots + |a_1| + |a_0|)x^k \\
&= ax^k.
\end{aligned}
$$

□

It is not hard to see that the following facts are true.

**Observation 1.9.2** *The "Order"-notations have the following properties:*

*(a)* $g(n) = O(f(n))$ *if and only if* $f(n) = \Omega(g(n))$.

*(b) The order of the sum of two functions is given by the order of the faster growing function:* $f(n) + g(n) = O(\max\{f(n), g(n)\})$.

*(c) The relation represented by "O" is transitive.*

*(d) For the logarithmic order* $O(\log n)$ *the base is irrelevant since* $\log_b n = \log_a n \cdot \log_b a$.

*(e) Exponential functions grow faster than polynomial functions:* $n^k = O(b^n)$ *for all* $k > 0$ *and* $b > 1$. *Conversely, logarithmic functions grow more slowly than polynomial functions.*

*(f)* $f(n) = \Theta(g(n))$ *if and only if* $g(n) = \Theta(f(n))$.

For our purpose we will use the following "classes of order", which are defined in terms of the input size $n$:

| Order | Name of the "class" | Remark |
|---:|---|---|
| $O(1)$ | constant | the function is bounded |
| $O(\log n)$ | logarithmic | the base is irrelevant |
| $O(n)$ | linear | |
| $O(n \log n)$ | log-linear | the base is irrelevant |
| $O(n^2)$ | quadratic | |
| $O(n^3)$ | cubic | |
| $\vdots$ | | |
| $O(n^k)$ | polynomial | $k$ is a fixed positive integer |

Mention that the previous table shows the ”slow growing” orders, this table the ”fast growing” ones:

| Order | Name of the ”class” | Remark |
|---|---|---|
| $O(c^n)$ | exponential | $c > 1$ is a fixed positive real number |
| $\vdots$ | | |
| $O(n!)$ | factorial | $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$ |
| $\vdots$ | | |
| $\Omega(2^{2^n})$ | superexponential | |

In particular, we say that the function $f$ is polynomially bounded if there is a positive integer $k$ such that $f(n) = O(n^k)$.

Often we have no exact formula for counting the number of combinatorial objects of some kind, but we can describe its asymptotic behavior. Then we use the following notation: Let $f$ and $g$ be functions from the positive integers into the real numbers, then

(i) The function $g(n)$ is said to be growing faster than $f(n)$, denoted $f(n) = o(g(n))$, if

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0. \tag{1.76}$$

(ii) The function $g(n)$ is said to be approximately $f(n)$, denoted $f(n) \approx g(n)$, if

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 1. \tag{1.77}$$

This notation allows us to concentrate on the dominating term in an expression describing a lower or upper bound and to ignore any multiplicative constants.

**Observation 1.9.3** *The asymptotic-notations have the following properties:*

*(a)* *The relation represented by ”o” is transitive.*
*In particular, we have the following increasing sequence of functions:*

$$c, \log\log n, \log n, n, n \cdot \log n, n^2, n^3, c^n, n!, n^n. \tag{1.78}$$

*(b)* *If $f(n) \approx g(n)$, then $f(n) = \Theta(g(n))$, but not vice versa.*

A broader discussion on the growth of functions can be found in [1].

### 1.9.2 The order of magnitude of the factorials

The quantity $n!$ (factorial) increases very quickly. Particularly $20! \approx 2.432 \cdot 10^{18}$. We can describe the order of growing by the following considerations. In calculus, an integral can be regarded as the area under a curve, and we can approximate this area by adding up long, "skinny" rectangles that touch the curve. Consider the function $\ln x$. Then

$$\sum_{k=1}^{n-1} \ln k \leq \int_1^n \ln x dx \leq \sum_{k=2}^{n} \ln k$$

$$\ln \prod_{k=1}^{n-1} k \leq n \ln n - n + 1 \leq \ln \prod_{k=2}^{n} k$$

$$\ln \frac{n!}{n} \leq \ln n^n - n + 1 \leq \ln n!$$

$$\frac{n!}{n} \leq e\frac{n^n}{e^n} \leq n!$$

which gives the

**Observation 1.9.4** *(Stirling's inequalities)*

$$e\frac{n^n}{e^n} \leq n! \leq en\frac{n^n}{e^n}. \tag{1.79}$$

The following approximation is essentially harder to prove and we will omit this calculation.

**Remark 1.9.5** *(Stirling's equality)*

$$n! \approx \sqrt{2\pi n} \cdot \frac{n^n}{e^n}. \tag{1.80}$$

# 2

# SELECTING OBJECTS

Consider a set of $n$ objects. How many ways are there of selecting $k$ from these? We will distinguish two kinds of choosing:

- ordered or unordered;
- repititions allowed or not.

This gives us four distinct questions.

## 2.1 THE NUMBER OF SUBSETS

As introductionary example choose two elements from $\{1, 2, 3, 4\}$, where we respect its order, but ignore repititions:

| | | |
|---|---|---|
| 1,2 | 1,3 | 1,4 |
| 2,1 | 2,3 | 2,4 |
| 3,1 | 3,2 | 3,4 |
| 4,1 | 4,2 | 4,3 |

More systematically, if we select only $k$ objects, we start with $n$ possibilities and count down $k$ numbers, the last one will be $n - k + 1$. Hence, we have the following theorem.

**Theorem 2.1.1** *The number of subsets with $k$ ordered elements of a set with $n$ elements is*

$$n(n-1)\cdots(n-k+1)=\frac{n!}{(n-k)!}. \tag{2.1}$$

From this we can easily derive one of the most important counting results.

**Theorem 2.1.2** *The number of subsets containing $k$ elements of a set with $n$ elements is*

$$\frac{n(n-1)\cdots(n-k+1)}{k!}=\frac{n!}{k!(n-k)!}. \tag{2.2}$$

*Proof.* In 2.1.1 we counted ordered subsets. If we want to know the number of unordered subsets, then every subset was counted exactly $k!$ times, namely all possible orderings of the elements. So we have to divide this number by $k!$ to get the assertion.

□

As an example choose two elements from $\{1,2,3,4\}$:

| | | |
|---|---|---|
| 1,2 | 1,3 | 1,4 |
| | 2,3 | 2,4 |
| | | 3,4 |

The number defined in 2.1.2 is such an important quantity that there is a special notation for it:

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}, \tag{2.3}$$

read "$n$ choose $k$". These numbers are also called binomial coefficients; we will later see why.[1] In view of 2.1.1 we will write $\binom{X}{k}$ for the collection of all subsets

[1] Of course, for a calculation of a simple binomial coefficients it is not pleasant to use this formulae; better:

$$\binom{n}{k}=\frac{n}{k}\cdot\frac{n-1}{k-1}\cdot\frac{n-2}{k-2}\cdots\frac{n-k+1}{1}.$$

of $X$ with exactly $k$ elements. This gives for the power set

$$\mathcal{P}(X) = \bigcup_{k=0}^{|X|} \binom{X}{k}. \tag{2.4}$$

Recall 1.2.2 which says that the power set $\mathcal{P}(X)$ contains $2^{|X|}$ elements.

## 2.2 BINOMIAL COEFFICIENTS

Binomial coefficients play a very important role in combinatorics. Consequently, we will investigate these quantities more extensively.

### 2.2.1 The Pascal triangle

We have two descriptions for the binomial coefficients:

**Arithmetic:** $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

**Set theoretic:** $\binom{n}{k}$ = number of choosing $k$ elements from $n$.

Consequently, we have two methods to find facts.

**Observation 2.2.1**

$$\binom{n}{k} = \binom{n}{n-k}. \tag{2.5}$$

*Proof.* Algebraic this is obvious.
Since selecting the $k$ elements out of $n$, we're in effect selecting the $n-k$ unchosen elements.

□

**Observation 2.2.2**

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}. \tag{2.6}$$

*Proof.*

$$\begin{aligned}\binom{n}{k}+\binom{n}{k-1} &= \frac{n!}{k!(n-k)!}+\frac{n!}{(k-1)!(n-k+1)!}\\ &= \frac{n!(n-k+1)}{k!(n-k+1)!}+\frac{n!k}{k!(n-k+1)!}\\ &= \frac{n!(n-k+1+k)}{k!(n-k+1)!}\\ &= \frac{(n+1)!}{k!(n+1-k)!}\\ &= \binom{n+1}{k}.\end{aligned}$$

Alternatively, consider $n+1$ objects $x_1,\ldots,x_{n+1}$. A choice of $k$ of the objects may or may not include $x_{n+1}$. If it does not, then $k$ objects have to be chosen from $x_1,\ldots,x_n$ and there are $\binom{n}{k}$ such choices. If it does contain $x_{n+1}$, then $k-1$ further objects have to be chosen from $x_1,\ldots,x_n$, and there are $\binom{n}{k-1}$ such choices. The result now follows from the addition principle.

□

A nice description of the binomial coefficients is given by the so-called Pascal's triangle, which displays 2.2.1 and 2.2.2:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| row 0 | | | | | | | 1 | | | | | | | |
| row 1 | | | | | | 1 | | 1 | | | | | | |
| row 2 | | | | | 1 | | 2 | | 1 | | | | | |
| row 3 | | | | 1 | | 3 | | 3 | | 1 | | | | |
| row 4 | | | 1 | | 4 | | 6 | | 4 | | 1 | | | |
| row 5 | | 1 | | 5 | | 10 | | 10 | | 5 | | 1 | | |
| row 6 | 1 | | 6 | | 15 | | 20 | | 15 | | 6 | | 1 | |
| ⋮ | | | | | | | | | | | | | | ⋱ |

In the $n$th row, the entries are the binomial coefficients $\binom{n}{k}$ for $k=0,\ldots,n$.

### 2.2.2 The binomial theorem

We called the numbers $\binom{n}{k}$ the binomial coefficients. Now we will see why. Consider

$$\begin{array}{rcl}(x+y)^0 & = & 1; \\ (x+y)^1 & = & x+y; \\ (x+y)^2 & = & x^2+2xy+y^2; \\ (x+y)^3 & = & x^3+3x^2y+3xy^2+y^3; \\ (x+y)^4 & = & x^4+4x^3y+6x^2y^2+4xy^3+y^4.\end{array}$$

Note that the coefficients are the entries of the Pascal' triangle. And indeed,

**Theorem 2.2.3** *(The binomial theorem) For any real numbers $x$ and $y$ and nonnegative integers $n$*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \tag{2.7}$$

*holds.*

*Proof.*

$$(x+y)^n = \underbrace{(x+y)\cdots(x+y)}_{n-\text{times}}. \tag{2.8}$$

So the coefficient of the term $x^{n-k}y^k$ is the number of ways of getting $x^{n-k}y^k$ when the $n$ brackets are multiplied out. Each term in the expansion is the product of one term from each bracket; so $x^{n-k}y^k$ is obtained as many times as we can choose $y$ from $k$ of the brackets and $x$ from the remaining $n-k$ brackets. But this can be done just in $\binom{n}{k}$ ways.

□

**Corollary 2.2.4** *For any real number $y$ and nonnegative integers $n$*

$$(1+y)^n = \sum_{k=0}^{n} \binom{n}{k} y^k \tag{2.9}$$

*holds.*

This corollary is the origin of several important facts. First the entries in each row of the Pascal triangle satisfy the following equations:

$$\sum_{k=0}^{n}\binom{n}{k} = 2^n. \tag{2.10}$$

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0. \tag{2.11}$$

Second, the following inequality holds.

**Observation 2.2.5** *(Bernoulli's inequality) Let $a$ be a positive real number and let $n \geq 2$ be an integer. Then*

$$(1+a)^n > 1 + na. \tag{2.12}$$

*Proof.* In view of 2.2.4 we have

$$\begin{aligned}(1+a)^n &= \binom{n}{0}a^0 + \binom{n}{1}a^1 + \underbrace{\binom{n}{2}a^2 + \ldots}_{>0}\\ &> 1 + na.\end{aligned}$$

□

## 2.2.3 The order of growing for binomial coefficients

To find orders of growing and asymptotic behaviour of the binomial coefficients we start with the so-called trinomial revision:

**Lemma 2.2.6**

$$\binom{n}{k}\binom{k}{m} = \binom{n}{m}\binom{n-m}{k-m}. \tag{2.13}$$

*Proof.* The left-hand side counts the ways to select a group of $k$ elements from a set of $n$ elements and then (multiplication principle) to select a subset of $m$

of this group.
Equivalently, counted on the right-hand side, we could first select the subset of $m$ elements from the set of $n$ elements and then select the remaining $k-m$ elements of the group from the remaining $n-m$ elements.

□

A specific application of 2.2.6 with $m=1$ is

**Corollary 2.2.7**

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}. \tag{2.14}$$

On one hand, repeated application of 2.2.7 gives us

$$\begin{aligned}\binom{n}{k} &= \frac{n}{k}\cdot\frac{n-1}{k-1}\cdot\frac{n-2}{k-2}\cdots\frac{n-k+1}{1}\\ &\geq \frac{n}{k}\cdots\frac{n}{k}\\ &= \left(\frac{n}{k}\right)^k. \end{aligned} \tag{2.15}$$

On the other hand

$$\begin{aligned}\binom{n}{k} &= \frac{n}{k}\cdot\frac{n-1}{k-1}\cdot\frac{n-2}{k-2}\cdots\frac{n-k+1}{1}\\ &\leq \frac{n^k}{k!}\\ &\leq \frac{1}{e}\left(\frac{en}{k}\right)^k, \end{aligned} \tag{2.16}$$

using the Stirling inequality. Hence, considering together

**Theorem 2.2.8**

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \frac{1}{e}\left(\frac{en}{k}\right)^k. \tag{2.17}$$

For the sum of binomial coefficients we know $2^n$ as upper bound. A partially better bound is given by the following formulae.

**Theorem 2.2.9** *([58]) Let $k \leq n$, then*

$$\sum_{j=0}^{k} \binom{n}{j} \leq \left(\frac{en}{k}\right)^k. \tag{2.18}$$

*Proof.* We use 2.2.4:

$$\binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \ldots + \binom{n}{n}x^n = (1+x)^n \tag{2.19}$$

for all real numbers $x$. In particular for $0 < x < 1$

$$\binom{n}{0} + \binom{n}{1}x + \ldots + \binom{n}{k}x^k \leq (1+x)^n. \tag{2.20}$$

Dividing this by $x^k$ we get

$$\frac{1}{x^k}\binom{n}{0} + \frac{1}{x^{k-1}}\binom{n}{1} + \ldots + \binom{n}{k} \leq \frac{(1+x)^n}{x^k}. \tag{2.21}$$

Since $x < 1$ we have

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{k} \leq \frac{(1+x)^n}{x^k}. \tag{2.22}$$

Note that the left-hand side is independent of the value of $x$; in particular we can use $x = k/n$ which gives us

$$\frac{(1+x)^n}{x^k} = \frac{\left(1+\frac{k}{n}\right)^n}{\left(\frac{k}{n}\right)^k} \leq \left(e^{\frac{k}{n}}\right)^n \left(\frac{n}{k}\right)^k = e^k \left(\frac{n}{k}\right)^k. \tag{2.23}$$

To get this result we have to apply some calculus which shows $1 + x \leq e^x$, so that we come to the inequality

$$\frac{\left(1+\frac{k}{n}\right)^n}{\left(\frac{k}{n}\right)^k} \leq \left(e^{\frac{k}{n}}\right)^n \left(\frac{n}{k}\right)^k.$$

Hence the assertion.

□

The bound in 2.2.9 is essentially less than $2^n$ when

$$k \leq e^{\sqrt{\ln 2} - 1} \cdot n = 0.84582\ldots \cdot n. \tag{2.24}$$

### 2.2.4 The bird's-eye view of Pascal's triangle

Let us ask a more quantitative question about the shape of a row in Pascal's triangle: What is the ratio of any binomial coefficients in a row to the largest b.c. in this row? It is not hard to see that the element in the middle of the row is the largest. To make it easier we consider the case when $n$ is even, and write $n = 2m$. Then the largest, middle entry in the $n$th row is $\binom{2m}{m}$. Now consider the binomial coefficient that is $t$ steps from the middle and compare it with the largest element, that means we are interested in the term $\binom{2m}{m-t}/\binom{2m}{m}$.

**Theorem 2.2.10**

$$e^{-\frac{t^2}{m-t+1}} \leq \frac{\binom{2m}{m-t}}{\binom{2m}{m}} \leq e^{-\frac{t^2}{m+t}}. \tag{2.25}$$

*Proof.* To derive the formula, we take the ratio to its reciprocal, which is larger than 1.

$$\begin{aligned}
\frac{\binom{2m}{m}}{\binom{2m}{m-t}} &= \frac{\frac{(2m)!}{m!m!}}{\frac{(2m)!}{(m-t)!(m+t)!}} \\
&= \frac{(m-t)!(m+t)!}{m!m!} \\
&= \frac{(m+t)(m+t-1)\cdots(m+1)}{m(m-1)\cdots(m-t+1)} \\
&= \frac{m+t}{m}\cdot\frac{m+t-1}{m-1}\cdots\frac{m+1}{m-t+1}.
\end{aligned}$$

to handle with this term the following idea helps: Take the logarithm.

$$\ln\left(\frac{m+t}{m}\right) + \ln\left(\frac{m+t-1}{m-1}\right) + \ldots + \ln\left(\frac{m+1}{m-t+1}\right).$$

We estimate the logarithms using the well-known fact

$$\frac{x-1}{x} \leq \ln x \leq x - 1. \tag{2.26}$$

For a typical term in the sum we have

$$\ln\left(\frac{m+t-k}{m-k}\right) \leq \frac{m+t-k}{m-k} - 1 = \frac{t}{m-k},$$

and so we get by

$$\begin{aligned}
&\ln\left(\frac{m+t}{m}\right)+\ln\left(\frac{m+t-1}{m-1}\right)+\ldots+\ln\left(\frac{m+1}{m-t+1}\right)\\
\le\ & \frac{t}{m}+\frac{t}{m-1}+\ldots+\frac{t}{m-t+1}\\
\le\ & \frac{t}{m-t+1}+\frac{t}{m-t+1}+\ldots+\frac{t}{m-t+1}\\
=\ & \frac{t^2}{m-t+1}
\end{aligned}$$

an upper bound on the logarithm of the ratio. To get an upper bound on the ratio itself, we have to apply the exponential function.
Similar the lower bound.

□

The lower and upper bounds in 2.2.10 are quite similar to the (imprecise) approximation

$$\frac{\binom{2m}{m-t}}{\binom{2m}{m}} \approx e^{-t^2/m}, \tag{2.27}$$

which is the famous Gauß curve.

### 2.2.5 Splits

Let $\mathcal{U}$ be a universe of $n$ elements. A pair $(S, S^c)$ of two nonempty subsets of $\mathcal{U}$ is called a split.
If we choose $k$ elements for $S$, $0 < k < n$, then we also choose $n-k$ elements for $S^c$. Hence, each selecting of $k$ elements creates a split. We can do it in $\binom{n}{k}$ ways. But we find each split twice, namely by choosing $k$, and by choosing $n-k$. Consequently, the number of splits is

$$\begin{aligned}
\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k} &= \frac{1}{2}\left(\sum_{k=0}^{n}\binom{n}{k}-\binom{n}{0}-\binom{n}{n}\right)\\
&= \frac{1}{2}(2^n-1-1)\\
&= 2^{n-1}-1.
\end{aligned}$$

**Theorem 2.2.11** *The number of splits of a set of $n$ elements is $2^{n-1}-1$.*

### 2.2.6 The multinomial theorem

In expanding $(x_1 + x_2 + x_3)^7$, we think of writing down seven $(x_1 + x_2 + x_3)$ terms in a row, and then adding up $x_1^i x_2^j x_3^k$ for all ways of selecting $x_1$ from $i$ of the terms, $x_2$ from $j$ of the terms and $x_3$ from $k$ of the terms. Note that $i + j + k = 7$ in each case. In view of 1.6.2 we get

**Theorem 2.2.12** *(The multinomial theorem) Let $x_1, \ldots, x_m$ be any real numbers and let $n$ be a nonnegative integer, then*

$$(x_1 + \ldots + x_m)^n = \sum_{k_1+\ldots+k_m=n} \frac{n!}{k_1! \cdots k_m!} x_1^{k_1} \cdots x_m^{k_m}, \tag{2.28}$$

*where we read the sum sign that appears in the formula as "the sum over all lists $k_1, \ldots, k_m$ such that $k_1 + \ldots + k_m = n$".*[2]

With $x_1 = \ldots = x_m = 1$ we get (again) 1.6.3.

## 2.3 THE HARDY-WEINBERG EQUILIBRIUM

A Mendelian population may be considered to be a group of sexually reproducing organisms with a relatively close degree of genetic relationships. If all the gametes produced by a Mendelian population are considered as a hypothetical mixture of genetic units from which the next generation will arise, we have the concept of a gene pool.

### 2.3.1 The Hardy-Weinberg law

Consider zygotes with two factors, one from each parent. Then the zygotes are called diploid. How does the "genetic make-up" of population change over generations? Suppose that what happens at a given locus is independent of what happens at any other, and focus on changes at a single locus. Furthermore, suppose that there are two and only two alleles $A$ and $B$ that may sit at

[2] The number of terms which are to sumarize is equal $\binom{m+n-1}{n}$. We will prove this later in 2.4.2.

this locus. A given individual may then have one of three genotypes: the homozygotes $AA$ or $BB$ or the heterozygote $AB$.

Let $p$ be the frequency of allele $A$ in a population, defined by

$$p = \frac{\text{number of allele } A}{\text{total number of alleles}}. \tag{2.29}$$

Similarly, let $q$ be the frequency of allele $B$. Of course, $p + q = 1$.
We are interested in the frequencies of the genotypes $AA$, $AB$ and $BB$, denoted by $x$,$y$ and $z$, respectively. Then, assuming that the population contains $n$ individuals, the allele $A$ can be found $x \cdot 2n + y \cdot n$ times. Hence,

$$p = \frac{x \cdot 2n + y \cdot n}{2n} = x + \frac{y}{2}. \tag{2.30}$$

Similarly, for allele $B$,

$$q = \frac{z \cdot 2n + y \cdot n}{2n} = z + \frac{y}{2}. \tag{2.31}$$

Let us now make the following assumptions:

(a) Expected sex ratio is independent of genotype.

(b) Mating is random.

(c) Fertility and survivorship are independent of genotype.

(d) There is no mutation or migration.

Then the frequency of genotypes for the subsequent generation is

| mating | frequency of matings | $AA$ | $AB$ | $BB$ |
|---|---|---|---|---|
| $AA \times AA$ | $x^2$ | $x^2$ | 0 | 0 |
| $AA \times AB$ | $xy$ | $\frac{xy}{2}$ | $\frac{xy}{2}$ | 0 |
| $AA \times BB$ | $xz$ | 0 | $xz$ | 0 |
| $AB \times AA$ | $xy$ | $\frac{xy}{2}$ | $\frac{xy}{2}$ | 0 |
| $AB \times AB$ | $y^2$ | $\frac{y^2}{4}$ | $\frac{y^2}{2}$ | $\frac{y^2}{4}$ |
| $AB \times BB$ | $yz$ | 0 | $\frac{yz}{2}$ | $\frac{yz}{2}$ |
| $BB \times AA$ | $xz$ | 0 | $xz$ | 0 |
| $BB \times AB$ | $yz$ | 0 | $\frac{yz}{2}$ | $\frac{yz}{2}$ |
| $BB \times BB$ | $z^2$ | 0 | 0 | $z^2$ |
| sum | | $x_1$ | $y_1$ | $z_1$. |

Then

$$x_1 = x^2 + xy + \frac{y^2}{4} = \left(x + \frac{y}{2}\right)^2 = p^2,$$

$$y_1 = xy + 2xz + \frac{y^2}{2} + yz = 2\left(x + \frac{y}{2}\right)\left(z + \frac{y}{2}\right) = 2pq,$$

and

$$z_1 = \frac{y^2}{4} + yz + z^2 = \left(\frac{y}{2} + z\right)^2 = q^2.$$

Furthermore, for the frequency of the alleles in the subsequent generation

$$p_1 = x_1 + \frac{y_1}{2} = p^2 + pq = p(p+q) = p,$$

and

$$q_1 = z_1 + \frac{y_1}{2} = q^2 + pq = q(p+q) = q.$$

It follows immediately that $p_{n+1} = p_n$ and $q_{n+1} = q_n$, where $n$ is the generation number. In other words, $p_n$ and $q_n$ are constants independent of $n$. Let us go back to calling them $p$ and $q$ again.

Generation 0 is known as the parental generation ($P = F_0$), and generation $n$ as the $n$th filial generation ($F_n$).

**Theorem 2.3.1** *(Hardy, Weinberg) Under the assumptions listed above*

*(a)* *Allele frequencies $p$ and $q$ remain unchanged from generation to generation, and are therefore the same in the filial generations as in the parental generation;*

*(b)* *From generation $F_1$ onwards*[3] *the genotype frequencies $x$, $y$ and $z$ remain unchanged, and are therefore the same in the filial generations as in the parental generation with $x = p^2$, $y = 2pq$ and $z = q^2$.*

An equilibrium has come to mean pretty much the same as stability that is a system which is largely unaffected by internal or external changes since it is easily returned to its original condition after being disturbed.

[3] but not necessarily for $F_0$

### 2.3.2 Selecting pressure

In absence of selection there is no evolution. Now we are interested in including this work.

Let the allele frequencies at the end of gametic phase of generation $n$ be $p_n$ and $q_n$, and the genotype frequencies $x_n$, $y_n$ and $z_n$. Then we saw that $x_n = p_n^2$, $y_n = 2p_nq_n$ and $z_n = q_n^2$. Now we introduce a selection pressure by the following considerations. Let the probability of survival from zygotic phase to breeding phase for the various genotypes be in the ratio $w_x : w_y : w_z$. These values measured the relative fitness of a genotype in terms of its reproductive success. Usually norming $w_x = 1$ the quantities $w_y$ and $w_z$ are the relative selection values of genotypes $AB$ and $BB$ relative to to the probability of survival of genotype $AA$. Then at the breeding phase the ratios of the genotypes $AA$, $AB$ and $BB$ have been modified to

$$w_xp_n^2 : 2w_yp_nq_n : w_zq_n^2, \tag{2.32}$$

so that allele frequencies are now in the ratio

$$w_xp_n^2 + w_yp_nq_n : w_yp_nq_n + w_zq_n^2. \tag{2.33}$$

Similar to our computations above we find the following equation of mathematical population genetics.

**Theorem 2.3.2** *(Fisher, Haldane, Wright) Under the assumptions listed above the allele frequencies follows from generation to generation by*

$$\begin{aligned} p_{n+1} &= \frac{(w_xp_n + w_yq_n)p_n}{w_xp_n^2 + 2w_yp_nq_n + w_zq_n^2} \quad \textit{and} & (2.34) \\ q_{n+1} &= \frac{(w_yq_n)q_n + w_yp_n}{w_xp_n^2 + 2w_yp_nq_n + w_zq_n^2}. & (2.35) \end{aligned}$$

The extant of change in the frequency of allele $B$ per generation is denoted by $\Delta q$. It is not hard to see that

$$\Delta q = q_{n+1} - q_n = \frac{pq(p(w_y - w_x) + q(w_z - w_y))}{w_xp^2 + 2w_ypq + w_zq^2}. \tag{2.36}$$

## 2.3.3 Dominance

We described the three genotypes assigned with the following fitness values and initial frequencies:

| genotype | $AA$ | $AB$ | $BB$ |
|---|---|---|---|
| fitness | $w_x$ | $w_y$ | $w_z$ |
| frequency | $p^2$ | $2pq$ | $q^2$ |

In the following, we assume that $A$ is the original allele in the population. We shall also assume that the population is diploid, and therefore the initial population consists only of one genotype, namely $AA$.
The fitness of the newly created genotypes $AB$ and $BB$ will depend on the mode of interaction between $A$ and $B$. In this sense we distinguish the following cases.

- In dominant selection the two homozygotes have different fitness values, whereas the fitness of the heterozygote is the same as the fitness of one of the homozygous genotypes.

  | genotype | $AA$ | $AB$ | $BB$ |
  |---|---|---|---|
  | fitness | 1 | $1+s$ | $1+s$ |

  From (2.36) we obtain the following change in the frequency of allele $B$ per generation:

  $$\Delta q = \frac{sp^2q}{1-s-p^2s}. \tag{2.37}$$

- The new allele is recessive

  | genotype | $AA$ | $AB$ | $BB$ |
  |---|---|---|---|
  | fitness | 1 | 1 | $1+s$ |

  From (2.36) we obtain the following change in the frequency of allele $B$ per generation:

  $$\Delta q = \frac{spq^2}{1+sq^2}. \tag{2.38}$$

- In codominant selection the two homozygotes have different fitness, whereas the fitness of the heterozygote is the mean of the fitness of the two homozygous genotypes.

| genotype | $AA$ | $AB$ | $BB$ |
|---|---|---|---|
| fitness | 1 | $1+s$ | $1+2s$ |

From (2.36) we obtain the following change in the frequency of allele $B$ per generation:

$$\Delta q = \frac{spq}{1+2spq+2sq^2}. \tag{2.39}$$

- In overdominant selection the heterozygote has the highest fitness.

| genotype | $AA$ | $AB$ | $BB$ |
|---|---|---|---|
| fitness | 1 | $1+s$ | $1+t$ |

where $s > 0$ and $s > t$.
From (2.36) we obtain the following change in the frequency of allele $B$ per generation:

$$\Delta q = \frac{pq(2sq-tq-s)}{1+2spq+tq^2}. \tag{2.40}$$

- In underdominant selection the heterozygote has the lowest fitness.

| genotype | $AA$ | $AB$ | $BB$ |
|---|---|---|---|
| fitness | 1 | $1+s$ | $1+t$ |

where $s > 0$ and $s < t$.
In this case the change in the frequency is also described by (2.40).

More information is given by [40].

## 2.4 SELECTIONS WITH REPETITIONS

Recall what we discussed until now for selecting $k$ objects from a set of $n$:

| | ordered | unordered |
|---|---|---|
| no repititions | $\frac{n!}{(n-k)!}$ | $\binom{n}{k}$ |
| repitions allowed | $n^k$ | ? |

The first row we complete discussed.
The first case in the second we discussed in the first chapter.[4] As an example choose ordere two elements from $\{1, 2, 3, 4\}$ where repititions are allowed:

| | | | |
|---|---|---|---|
| 1,1 | 1,2 | 1,3 | 1,4 |
| 2,1 | 2,2 | 2,3 | 2,4 |
| 3,1 | 3,2 | 3,3 | 3,4 |
| 4,1 | 4,2 | 4,3 | 4,4 |

"?" in the second row means that we have to determine the number of ways these are to choose $k$ objects from $n$, where repititions are allowed, but where order does not matter.
As an example choose two elements from $\{1, 2, 3, 4\}$:

| | | | |
|---|---|---|---|
| 1,1 | 1,2 | 1,3 | 1,4 |
| | 2,2 | 2,3 | 2,4 |
| | | 3,3 | 3,4 |
| | | | 4,4 |

More systematically

**Theorem 2.4.1** *The number of unordered choices of $k$ from $n$, with repititions allowed is*

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}. \tag{2.41}$$

*Proof.* Any choice will consist of $x_1$ choices of the first object, $x_2$ choices of the second object, and so on, where the condition $x_1 + \ldots + x_n = k$ is satiesfied. We can represent such collection $x_1, \ldots, x_n$ of integers by a binary sequence:

$$\underbrace{0, \ldots, 0}_{x_1-\text{times}}, 1, \underbrace{0, \ldots, 0}_{x_2-\text{times}}, 1, \underbrace{0, \ldots, 0}_{x_3-\text{times}}, 1, \ldots, 1, \underbrace{0, \ldots, 0}_{x_n-\text{times}} \tag{2.42}$$

In this representation there will be $n-1$ times the digit 1 and $k$ times the digit 0, and so each sequence will be of length $n+k-1$, containing exactly $k$ 0s. Conversely, any such sequence corresponds to a nonnegative integer solution of $x_1 + \ldots + x_n = k$.
The $k$ 0s can be in any of the $n+k-1$ positions, so the number of such sequences is $\binom{n+k-1}{k}$.

[4] Do you remember where?

□

From the first fact in the proof we get

**Theorem 2.4.2** *The number of solutions for the equation $x_1 + \ldots + x_n = k$ in nonnegative integers $x_i$ equals*

$$\binom{n+k-1}{k}. \tag{2.43}$$

As an example consider

$$\begin{aligned} x + y + z &= 8 \\ \text{subject to} \quad x &\geq 2 \\ y &\geq 4 \end{aligned}$$

We substitute $x = 2 + u$, $y = 4 + v$ and solve $u + v + z = 2$ with help of 2.4.2 to find $\binom{3+2-1}{2} = 6$ solutions.

# 3

# PARTITIONS

Recall 1.6.2. We can prove the theorem in another way using the following considerations: If there are $n$ objects of $k$ types with $n_i$ of the $i$th type, $i = 1, \ldots, k$, where $n_1 + \ldots + n_k = n$, then the number of arrangements are

$$\begin{aligned} & \binom{n}{n_1} \cdot \binom{n-n_1}{n_2} \cdots \binom{n - \sum_{i=1}^{k-1} n_i}{n_k} \\ = \quad & \frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdots \frac{(n - \sum_{i=1}^{k-1} n_i)!}{n_k!} \\ = \quad & \frac{n!}{n_1! \cdot n_2! \cdots n_k!}. \end{aligned}$$

The quantity

$$\frac{n!}{n_1! \cdot n_2! \cdots n_k!} \tag{3.1}$$

is called the multinomial coefficient, see 2.2.12.

## 3.1 PARTITION OF A SET

A partition of a (not necessarily finite) set $S$ is a collection of subsets $S_i$, $i \in I$ of $S$ such that

(i) $S_i \neq \emptyset$ for all $i \in I$;

(ii) $S_i \cap S_j = \emptyset$ for $i \neq j$; and

(iii) $\bigcup_{i \in I} S_i = S$.

The subsets $S_i$ are called the parts of the partition. A partition of a set into exactly two parts is called a bipartition or a split.

In particular we will discuss classifications, which are hierarchies of partitions, and of great relevance in biology. In the book *The System of Nature* Linnaeus introduced a system still in use today. He gave every species two Latinized names; the first for the group it belongs to, the genus; and the second for the particular organism itself. Today we divide life into: Domain[1]; Kingdom; Phylum; Class; Order; Family; Genus; and Species.
More or less all these groups are artificial, insofar as their members are categorized according to agreed-upon levels of similarity rather than precise definitions. The exception is species, which are defined as a group of individual organisms that are able to interbreed and produce fertile offspring. This definition creates an equivalence relation, and the species are the equivalence classes. Each group is a partition of the set of all species. For example

| group \ species | human | fruit fly |
|---|---|---|
| Domain | Eukarya | Eukarya |
| Kingdom | Animalia | Animalia |
| Phylum | Chordata | Arthropoda |
| Class | Mammalia | Insecta |
| Order | Primata | Diptera |
| Family | Hominidae | Drosophilidae |
| Genus | Homo | Drosophila |
| Species | *sapiens* | *melanogaster* |

Remember 1.3.1 which declares a one-to-one correspondence between partitions and equivalence relations. A direct consequence of the multiplication principle is

**Observation 3.1.1** *If $\sim$ is an equivalence relation on a set $S$ with $n$ elements and each equivalence class has the same number $m$ of elements. Then $\sim$ has $n/m$ equivalence classes.*

[1] There are three domains. The first two, Bacteria and Archea, are made up of many microscopic single-celled organisms. The third domain, Eukarya, is diverse.

## 3.2 PARTITIONS OF A GIVEN SIZE

Now we are interested in the number of partitions with specified, but not necessarily equal, part sizes. Consider a set $S$ of $n$ elements and a partition of $S$ into $\alpha_1$ parts of size 1, $\alpha_2$ parts of size 2, up to $\alpha_n$ parts of size $n$, where, of course, $1 \leq i \leq n$ and

$$\sum_{i=1}^{n} i\alpha_i = n. \tag{3.2}$$

Such a partition is called of the type $[1]\alpha_1[2]\alpha_2 \ldots [n]\alpha_n$. Recall 1.6.2. The $n$ elements can be placed in $n!$ ways. To count distinct partitions we have to take into account the ways of ordering the elements within the parts and the ways of ordering the parts of the same size $i$. Hence,

**Theorem 3.2.1** *The number of partitions of type* $[1]\alpha_1[2]\alpha_2 \ldots [n]\alpha_n$ *is*

$$\frac{n!}{\prod_{i=1}^{n}(i!)^{\alpha_i} \cdot \alpha_i!}. \tag{3.3}$$

In particular, the number of partitions of type $[2]m$, with $m = \frac{n}{2}$ is

$$\frac{n!}{2^m \cdot m!}. \tag{3.4}$$

## 3.3 THE STIRLING NUMBERS OF THE FIRST KIND

We are interested to calculate the number of partitions. In a first view we consider the similar question that the elements in the parts are ordered, that means we consider cycles in permutations.

The Stirling numbers of the first kind is defined as follows: $s(n,k)$ is the number of permutations of $1,\ldots,n$ consisting of exactly $k$ cycles.

As an example consider the permutations of the set $\{1,2,3,4\}$. Clearly, these 24 permutations can be classified according to the cycles they have, as follows.

1. There are 6 permutations with exactly one cycle: (1234), (1324), (2134), (2314), (3124) and (3214).
2. There are 11 permutations with exactly two cycles: (123)(4), (124)(3), (134)(2), (234)(1), (132)(4), (142)(3), (143)(2), (243)(1), (12)(34), (13)(24) and (14)(23).
3. There are 6 permutations with exactly three cycles: (12)(3)(4), (13)(2)(4), (14)(2)(3), (23)(1)(4), (24)(1)(3) and (34)(1)(2).
4. There is only one permutation with four cycles: (1)(2)(3)(4).

Here are several elementary facts about the Stirling number of the first kind.[2] Obviously, $s(n,n) = 1$ and

$$\sum_{k=1}^{n} s(n,k) = n!. \tag{3.5}$$

**Theorem 3.3.1**

$$s(n,1) = (n-1)!. \tag{3.6}$$

*Proof.* If we select the first $n-1$ elements, then the $n$th is determined. Selecting $n-1$ elements can be done in $(n-1)!$ ways.

□

**Theorem 3.3.2**

$$s(n,n-1) = \binom{n}{2}. \tag{3.7}$$

The proof remains as an exercise for the reader.

**Theorem 3.3.3** *For $n, k \geq 2$ it holds*

$$s(n,k) = (n-1) \cdot s(n-1,k) + s(n-1,k-1). \tag{3.8}$$

[2] Further results we will find later when considering the Stirling number of the second kind.

*Proof.* Consider $n$ elements and the $n$th element explicitely. $n$ either forms a 1-cycle on its own or can be slotted into a cycle in one of the $s(n-1, k-1)$ permutations of $1, \ldots, n-1$.

□

It is customary to write the Stirling numbers as Stirling's triangle (of the first kind) in the form of a right triangle.

| $n \setminus k$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1 | | | | | |
| 2 | 1 | 1 | | | | |
| 3 | 2 | 3 | 1 | | | |
| 4 | 6 | 11 | 6 | 1 | | |
| 5 | 24 | 50 | 35 | 10 | 1 | |
| 6 | 120 | 274 | 225 | 85 | 15 | 1 |

## 3.4 THE STIRLING NUMBERS OF THE SECOND KIND

The Stirling number $S(n, k)$ of the second kind denotes the number of ways of partitioning of a set of $n$ elements into exactly $k$ parts.

As an example consider the partitions of the set $\{1, 2, 3, 4\}$. These partitions can be classified according to the number of parts they have, as follows.

1. There is only one partition with exactly one part: $\{\{1, 2, 3, 4\}\}$.
2. There are 7 partitions with exactly two parts:

   $$\begin{array}{llll} \{\{1,2,3\},\{4\}\} & \{\{1,2,4\},\{3\}\} & \{\{1,3,4\},\{2\}\} & \{\{2,3,4\},\{1\}\} \\ \{\{1,2\},\{3,4\}\} & \{\{1,3\},\{2,4\}\} & \{\{1,4\},\{2,3\}\} & \end{array}$$

3. There are 6 partitions with exactly three parts:

   $$\begin{array}{lll} \{\{1,2\},\{3\},\{4\}\} & \{\{1,3\},\{2\},\{4\}\} & \{\{1,4\},\{2\},\{3\}\} \\ \{\{2,3\},\{1\},\{4\}\} & \{\{2,4\},\{1\},\{3\}\} & \{\{3,4\},\{1\},\{2\}\} \end{array}$$

4. There is only one partition with four parts: $\{\{1\},\{2\},\{3\},\{4\}\}$.

Of course, for all $n \geq 1$, $S(n,1) = S(n,n) = 1$.

**Theorem 3.4.1** *For all $n \geq 2$,*

*(a)*

$$S(n,2) = 2^{n-1} - 1. \tag{3.9}$$

*(b)*

$$S(n,n-1) = \binom{n}{2}. \tag{3.10}$$

*Proof.* (a): $S(n,2)$ is the number of splits, which we counted in 2.2.11.
(b): On of the parts must have two elements. You can choose this part in $\binom{n}{2}$ ways.

□

**Theorem 3.4.2** *Whenever $1 < k < n$,*

$$S(n,k) = S(n-1,k-1) + k \cdot S(n-1,k). \tag{3.11}$$

*Proof.* Consider a partition of $\{1,\dots,n\}$ into $k$ parts; consider the element $n$.
Case 1: $n$ appears by itself as a 1-element part.
Then the remaining $n-1$ elements have to form a partition of $\{1,\dots,n-1\}$ into $k-1$ subsets. There are $S(n-1,k-1)$ ways in which this can be done.
Case 2: $n$ is in a part of size at least two.
Then we can think of partitioning $\{1,\dots,n-1\}$ into $k$ sets (which can be done in $S(n-1,k)$ ways) and then of adding $n$ in one of the $k$ sets (there are $k$ ways of doing this).
The addition and multiplication principles give the assertion.

□

There is an explicit formulae for the Stirling number of the second kind, but we will omit the proof[3]:

$$S(n,k) = \frac{1}{k!}\sum_{i=0}^{k-1}(-1)^i\binom{k}{i}(k-i)^n. \tag{3.12}$$

It is customary to write the Stirling numbers as Stirling's triangle in the form of a right triangle, where the second column is the number of splits and the diagonal under the main diagonal are the binomial coefficients.

| $n \setminus k$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1 | | | | | |
| 2 | 1 | 1 | | | | |
| 3 | 1 | 3 | 1 | | | |
| 4 | 1 | 7 | 6 | 1 | | |
| 5 | 1 | 15 | 25 | 10 | 1 | |
| 6 | 1 | 31 | 90 | 65 | 15 | 1 |

## 3.5 THE INTERRELATION BETWEEN THE STIRLING NUMBERS AND THE BELL NUMBERS

The Stirling number of the first kind must be at least as large as the number of the second kind, since vvery partition into nonempty parts leads to at least one arrangement of cycles:

**Theorem 3.5.1** *For all integers n and k it holds*

$$s(n,k) \geq S(n,k). \tag{3.13}$$

Equality holds in 3.5.1 when all the cycles are necessarily singletons or doubletons, because cycles are equivalent to subsets in such cases:

$$\begin{aligned} s(n,n) = S(n,n) &= 1, & (3.14)\\ s(n,n-1) = S(n,n-1) &= \binom{n}{2}. & (3.15)\end{aligned}$$

[3] Compare [3].

In all other cases the inequality in 3.5.1 is strict. Moreover, $s(n,1) = (n-1)! \gg 1 = S(n,1)$.

$B(n)$ is the total number of partitions of a set of $n$ elements, and is called a the Bell number:

$$B(n) = \sum_{k=1}^{n} S(n,k), \tag{3.16}$$

where we define $B(0) = 1 = S(0,0)$. Immediately from 3.5.1 we have

$$B(n) \leq n!. \tag{3.17}$$

**Theorem 3.5.2** *For all $n \geq 1$,*

$$B(n) = \sum_{k=0}^{n-1} \binom{n-1}{k} B(k). \tag{3.18}$$

*Proof.* Consider the $n$th element of a set which is partitioned. It is in one of parts of the partition with $j \geq 0$ other elements. There are $\binom{n-1}{j}$ ways of choosing these $j$ elements. The remaining $n-1-j$ elements can be partitioned in $B(n-1-j)$ ways. Hence,

$$\begin{aligned} B(n) &= \sum_{j=0}^{n-1} \binom{n-1}{j} B(n-1-j) \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} B(k), \end{aligned}$$

putting $n-1-j = k$.

□

3.5.2 yields an elegant (recursive) procedure for the computation of these numbers.

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $B(n)$ | 1 | 2 | 5 | 15 | 52 | 203 | 877 | 4140 |

# 4

# RECURRENCE RELATIONS

Recursion is a process that wraps back on itself and feeds the output of a process back in as the input.

Remember that $n!$ satisfies the following equations:

$$\begin{aligned} n! &= n \cdot (n-1)! \text{ for } n \geq 1; && (4.1)\\ 0! &= 1. && (4.2) \end{aligned}$$

Recall the triangles of the binomial coefficients $b(n,k) = \binom{n}{k}$, the Stirling numbers $s(n,k)$ of the first and $S(n,k)$ of the second kind. We constructed these triangles by

$$\begin{aligned} b(n,k) &= b(n-1,k) + b(n-1,k-1), && (4.3)\\ s(n,k) &= (n-1) \cdot s(n-1,k) + s(n-1,k-1), && (4.4)\\ S(n,k) &= k \cdot S(n-1,k) + S(n-1,k-1), && (4.5) \end{aligned}$$

compare 2.2.2, 3.3.3 and 3.4.2.

It often happens that in studying a sequence of numbers, a connection between the current value and several of the previous values is obtained. This connection is called a recurrence relation. We will discuss how such recurrences arise and how they may be solved. A general solution method is unknown, but we will find solutions for several specific and important cases.

In any case we have to handle recurrences with care: Consider the following sequence

$$a_n = \begin{cases} 1 + a_{\frac{n}{2}} & : \text{ n even} \\ 1 + a_{3n-1} & : \text{ otherwise} \end{cases}$$

with $a_1 = 1$. We get

$$\begin{aligned} a_5 &= 1 + a_{14} \\ &= 1 + (1 + a_7) = 2 + a_7 \\ &= 2 + (1 + a_{20}) = 3 + a_{20} \\ &= 3 + (1 + a_{10}) = 4 + a_{10} \\ &= 4 + (1 + a_5) = 5 + a_5, \end{aligned}$$

which means that $a_5$ is not defined.

## 4.1 LINEAR RECURRENCE RELATIONS

Let $k$ be a positive integer and let $c_1, \ldots, c_k$ be real numbers with $c_k \neq 0$. Then

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k} \tag{4.6}$$

with given real numbers

$$a_0, \ldots, a_{k-1}, \tag{4.7}$$

is called a linear recurrence relation (with constant coefficients) of order $k$. The equation (4.6) is called the recurrence equation, and the values (4.7) are called the initial conditions. Note that the order of the recurrence relation is, on the one hand, the number of the dependences on previous elements of the sequence, and on the other hand, the number of initial conditions.

**Observation 4.1.1** *If the sequences $a_n$ and $a'_n$ satisfy the recurrence equation (4.6), then the linear combination $\alpha \cdot a_n + \alpha' \cdot a'_n$ satisfies this as well.*

### 4.1.1 First order

Suppose that the population of a colony of ants doubles in each successive year. A colony is established with an initial population of $a_0 = a$ ants. How many ants will this colony have after $n$ years? Let $a_n$ denote this number. Then

$$a_n = 2 \cdot a_{n-1} = 2^2 \cdot a_{n-2} = \ldots = 2^{n-1} \cdot a_1 = 2^n \cdot a. \tag{4.8}$$

Another example is when we consider the well-known Towers of Hanoi: Consider $n$ discs, all of different sizes. A collection of discs forms a tower if they ordered according to their sizes, the largest at the bottom. Now

(i) All the $n$ discs form a tower on position 1.

(ii) There are two other positions: 2 and 3.

(iii) Move the original tower from position 1 to position 3 by moving only exactly one disc in one step, and by producing only towers at the positions.

What is minimum number of steps required? Let $a_n$ be the smallest number of steps required to move the $n$ discs. It is easy to see that: $a_1 = 1$, $a_2 = 3$ and $a_3 = 7$ (exercise).

What about $a_n$? Forget the bottom disc and move the remaining $n-1$ discs to position 2. To get this stage, $a_{n-1}$ steps are needed. Then move the disc from position 1 to position 3: one step. Now move the tower from position 2 to position 3. Altogether we need $a_{n-1} + 1 + a_{n-1}$ steps. Hence we have to solve

$$a_n = 2 \cdot a_{n-1} + 1, \tag{4.9}$$

with $a_1 = 1$.[1] Then

$$a_n = 2^n - 1. \tag{4.10}$$

This is an immediate consequence of the following formulae.

**Observation 4.1.2** *Let*

$$a_n = ca_{n-1} + g, \tag{4.11}$$

$n \geq 1$, *with given constants* $c$ *and* $g$, *and an initial condition*

$$a_0, \tag{4.12}$$

*be a (nonhomogeneous) recurrence relation of the first order. Then*

$$a_n = \begin{cases} c^n a_0 + \frac{c^n - 1}{c-1} g & : \quad c \neq 1 \\ a_0 + ng & : \quad c = 1 \end{cases}$$

*For* $0 < c < 1$ *we have*

$$a_n \to \frac{1}{1-c} g.$$

*Proof.*

$$a_n \quad = \quad c \cdot a_{n-1} + g$$

[1] In reality, we only proved $a_n \leq 2 \cdot a_{n-1} + 1$, since these moves suffice. As exercise show that so many moves are also necessary.

$$
\begin{aligned}
&= c \cdot (c \cdot a_{n-2} + g) + g = c^2 \cdot a_{n-2} + cg + g \\
&= c^2 \cdot (c \cdot a_{n-3} + g) + cg + g = c^3 \cdot a_{n-3} + c^2 g + cg + g \\
&\vdots \\
&= c^{n-1} \cdot a_1 + c^{n-2} g + \ldots + c^2 g + cg + g \\
&= c^n a_0 + c^{n-1} g + \ldots + c^2 g + cg + g \\
&= c^n a_0 + (c^{n-1} + \ldots + c^2 + c + 1) \cdot g.
\end{aligned}
$$

Now, we distinguish between $c = 1$ and $c \neq 1$.
For $0 < c < 1$ we have an convergent geometric sequence.

□

As exercise

(a) Discuss the equilibrium case.

(b) Consider the case $c = -1$, $g = 1$ and $a_0 = 1$.

### 4.1.2 Second order

In his famous book *Liber Abaci*, Fibonacci raised the following question

> A certain man put a pair of rabbits in a place surround on all sides by a wall. How many pairs of rabbits can be produced from that pair in a year if it is supposed that every month each pair begets a new pair which from the second month on becomes productive.

For convenience, we will count the rabbits in male-female pair. $F_0$ represents the initial population, and $F_i$ represents the population in the $i$th month.

$$f_i = |F_i| \tag{4.13}$$

denotes the total number of pairs in the $i$th generation.

| generation | $F_0$ | $F_1$ | $F_2$ | $F_3$ | $F_4$ | $F_5$ | $F_6$ |
|---|---|---|---|---|---|---|---|
| number of mature pairs | 0 | 1 | 1 | 2 | 3 | 5 | 8 |
| number of baby pairs | 1 | 0 | 1 | 1 | 2 | 3 | 5 |
| $f_i$ | 1 | 1 | 2 | 3 | 5 | 8 | 13 |

We can see from this table that

$$f_n = f_{n-1} + f_{n-2}, \tag{4.14}$$

for $n \geq 2$ with $f_0 = 1$ and $f_1 = 1$.

Now, we concentrate on the recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} \tag{4.15}$$

with given $a_0, a_1$, and constants $c_1, c_2$, where $c_2 \neq 0$. There is a very neat method of solving such relations. Substituting $\alpha^n$ with $\alpha \neq 0$ for $a_n$ in (4.15) gives $\alpha^n = c_1\alpha^{n-1} + c_2\alpha^{n-2}$, that is $\alpha^2 = c_1\alpha + c_2$. Consequently, $\alpha^n$ is a solution of (4.15) if and only if $\alpha$ is a solution of the so-called characteristic equation

$$x^2 = c_1 x + c_2. \tag{4.16}$$

We have to distinguish between two cases:

(i) We assume that $\alpha$ and $\beta$ are distinct solutions of (4.16). Then $\alpha^n$ and $\beta^n$ satisfy (4.15), hence its linear combination also:

$$a_n = d_1\alpha^n + d_2\beta^n. \tag{4.17}$$

Choose $d_1, d_2$ so that

$$\begin{aligned} a_0 &= d_1 + d_2 \\ a_1 &= d_1\alpha + d_2\beta. \end{aligned}$$

(ii) On the other hand, when the characteristic equation has a repeated root $\alpha$, that means of multiplicity two, then

$$x^2 - c_1 x - c_2 = (x - \alpha)^2 = x^2 - 2\alpha x + \alpha^2 \tag{4.18}$$

so that $c_1 = 2\alpha$ and $c_2 = \alpha^2$. Following $n\alpha^n$ also satisfies (4.15), since

$$\begin{aligned} c_1 a_{n-1} + c_2 a_{n-2} &= c_1(n-1)\alpha^{n-1} + c_2(n-2)\alpha^{n-2} \\ &= 2(n-1)\alpha^n - (n-2)\alpha^n \\ &= n\alpha^n \\ &= a_n. \end{aligned}$$

Choose $d_1, d_2$ so that

$$\begin{aligned} a_0 &= d_1 \\ a_1 &= d_1\alpha + d_2\alpha. \end{aligned}$$

Solving these equations we get the following theorem.

**Theorem 4.1.3** *Suppose that $\{a_n\}$ satisfies the recurrence relation*

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} \tag{4.19}$$

*with given $a_0, a_1$.*
*Let $\alpha$ and $\beta$ be the roots of the characteristic equation*

$$x^2 - c_1 x - c_2 = 0. \tag{4.20}$$

*(a) If $\alpha \neq \beta$ then*

$$a_n = d_1\alpha^n + d_2\beta^n, \tag{4.21}$$

*with*

$$d_1 = \frac{a_1 - a_0\beta}{\alpha - \beta}, \tag{4.22}$$

$$d_2 = \frac{a_0\alpha - a_1}{\alpha - \beta}. \tag{4.23}$$

*(b) If $\alpha = \beta$ then*

$$a_n = (a_0 + nd_2)\alpha^n, \tag{4.24}$$

*with*

$$d_2 = \frac{a_1}{\alpha} - a_0. \tag{4.25}$$

Recall Fibonacci's rabbits to find the characteristic equation in

$$x^2 - x - 1 = 0,$$

with the roots

$$\alpha, \beta = \frac{1 \pm \sqrt{5}}{2},$$

**Corollary 4.1.4** *The Fibonacci sequence*

$$f_n = f_{n-1} + f_{n-2}, \tag{4.26}$$

*for* $n \geq 2$ *with* $f_0 = 1$ *and* $f_1 = 1$*, has the solution*

$$f_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1+\sqrt{5}}{2} \right)^{n+1} - \left( \frac{1-\sqrt{5}}{2} \right)^{n+1} \right). \tag{4.27}$$

This result is strange, since $f_n$ is in any case an integer.[2]

In the nineteenth century Fibonnacci numbers were discovered in many natural forms. For example, many types of flower have a Fibonacci number of petals: certain types of daises tend to have 34 or 55 petals, while sunflowers have 89 or 144. The understanding of these relations are called phylotaxis, compare [24].

### 4.1.3 A general solution method

Generalising the method used in 4.1.3 we outline the theory for solving recurrence relations of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k} \tag{4.29}$$

with given

$$a_0, \ldots, a_{k-1}, \tag{4.30}$$

where the $c_i$'s are given constants, $c_k \neq 0$.
The recurrence allows us to compute $a_n$ for any $n$ we like, but it only gives indirect information. A solution to the recurrence in a ”closed form” helps us to understand what $a_n$ really stands for.

**Algorithm 4.1.5** *Let a recurrence relation (4.29), (4.30) be given. Then*

1. *Solve the characteristic equation*

$$x^k - c_1 x^{k-1} - c_2 x^{k-2} - \ldots - c_k = 0. \tag{4.31}$$

[2] The number

$$\frac{1+\sqrt{5}}{2} = 1.61803\ldots \tag{4.28}$$

is important in many parts of mathematics as well as in the art world since ancient times. Therefore is has a special name, the golden ratio.

*It has $k$ roots, some of which may be multiple.*[3]

2. *If $\alpha_1, \ldots, \alpha_r$ are the roots of (4.31), then $a_n = \alpha_i^n$ is a solution of the recurrence equation (4.29). Any linear combination of such solutions is also a solution, see 4.1.1.*
   *Compose a linear combination for the roots in the following sense: If root $\alpha$ has multiplicity $m$ then use*
   $$\alpha^n, n\alpha^n, n^2\alpha^n, \ldots, n^{m-1}\alpha^n. \tag{4.32}$$

3. *We need to be given the initial conditions of the first $k$ values (4.30).*
   *The $k$ equations can be solved if we insert these conditions. This forms a system of $k$ linear equations with $k$ unknowns, which is simple to solve.*

An example: $a_n$ satisfies the relation

$$a_n = -2a_{n-2} - a_{n-4}$$

with $a_0 = 0, a_1 = 1, a_2 = 2$ and $a_3 = 3$.
The characteristic equation is

$$0 = x^4 + 2x^2 + 1 = (x^2+1)^2.$$

The roots of this equation are $i$ and $-i$ (where $i = \sqrt{-1}$) and each root has multiplicity 2. so the general solution is

$$a_n = x_1 i^n + x_2 n i^n + x_3(-i)^n + x_4 n(-i)^n.$$

Substituting the initial conditions and solving the four simultaneous equations in the four unknowns $x_1, \ldots, x_4$ we obtain

$$a_n = -\frac{3}{2}i^{n+1} + (-\frac{1}{2} + i)ni^n + \frac{3}{2}i(-i)^n + (-\frac{1}{2} - i)n(-i)^n.$$

### 4.1.4 Nonhomogeneous recurrence relations

Consider recurrence relations of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k} + g(n) \tag{4.33}$$

[3] A solution is explicetely possible if $k \leq 4$, but maybe not for the case $k > 4$. For considerations behind these questions, the so-called Galois theory, see [81].

with given

$$a_0, \ldots, a_{k-1}, \tag{4.34}$$

where the $c_i$'s are given constants, $c_k \neq 0$, and $g(n)$ is a given function.
If $g(n)$ is zero, the recurrence relation is called homogeneous, which is what we discussed extensively above, and otherwise, nonhomogeneous.
A solution method uses the following fact, which is similar to 4.1.1.

**Observation 4.1.6** *The difference of two solutions of a nonhomogeneous recurrence equation is a solution of its homogeneous part.*

As an example consider the special case

$$a_n = a_{n-1} + g(n). \tag{4.35}$$

We have

$$\begin{aligned}
a_1 &= a_0 + g(1) \\
a_2 &= a_1 + g(2) = a_0 + g(1) + g(2) \\
a_3 &= a_2 + g(3) = a_0 + g(1) + g(2) + g(3) \\
&\vdots \\
a_n &= a_0 + g(1) + g(2) + \ldots + g(n) \\
a_n &= a_0 + \sum_{i=1}^{n} g(i),
\end{aligned}$$

which shows that we can solve this type of relation in terms of $n$, if we can find a suitable summation formula for

$$\sum_{i=1}^{n} g(i). \tag{4.36}$$

In particular the recurrence equation (4.35) with $g(n) = cn^2$ has the solution

$$a_n = a_0 + c \cdot \frac{n(n+1)(2n+1)}{6}. \tag{4.37}$$

For practice prove the following generalization.

**Theorem 4.1.7** *Consider the nonhomogeneous recurrence equation*

$$a_n + c_{n-1}a_{n-1} = cr^n, \tag{4.38}$$

*of first order, where $c$ is a constant.*

*(i) If $r^n$ is not a solution of the associated homogeneous equation*

$$a_n + c_{n-1}a_{n-1} = 0, \tag{4.39}$$

*then*

$$a_n = Ar^n \tag{4.40}$$

*is a solution of (4.38), where $A$ is a constant.*

*(ii) When $r^n$ is a solution of the associated homogeneous equation (4.39) then*

$$a_n = Anr^n \tag{4.41}$$

*is a solution of (4.38), where $A$ is a constant.*

## 4.2 THE DYNAMICS OF POPULATION GROWTH

Recall the ants colony or Fibonacci's rabbits. The growth of a population is a dynamical process, meaning that it represents a situation that changes over time.
The ebb and the flow of a population over time can be conveniently thought of as a list of numbers called the population sequence. Every population start with a initial population $N_0$, the 0th generation, and continues with $N_1, N_2, \ldots$, where $N_n$ is the size of the $n$th generation.

Graham et al. [39] discussed that "bee trees" provide a good and more realistic example of such a point of view. Let's consider the pedigree of a male bee. Each male (called a drone) is produced asexually from a female (called a queen); each female, however, has two parents, a male and a female.
The drone has one grandfather and one grandmother; he has one great-grandfather and two great-grandmothers, he has two great-great-grandfathers and three great-great-grandmothers. In general, by induction, we see that he has exactly $f_{n+1}$ great$^n$-grandfathers and $f_{n+2}$ great$^n$-grandmothers.

### 4.2.1 The Malthusian equation

Let us assume that the size $N_n$ of a population at time $n$ completely determines its size at time $n + 1$.[4] More formally:

$$N_{n+1} = f(N_n). \tag{4.42}$$

As an introductory example consider a continuous population censused at intervals. Let the probability of any given individual dying between censuses be $d$, and let the average number of births to any given individual in the same period be $b$. Then the total number of deaths is $d \cdot N_n$, the total number of births $b \cdot N_n$, and consequently

$$N_{n+1} = (1 + b - d) \cdot N_n. \tag{4.43}$$

Such a linear model is known as the Malthusian equation. The parameter $\lambda = 1 + b - d$ is called the growth ratio. For

(a) $\lambda < 1$, we have decay;

(b) $\lambda > 1$, we have actual growth; and

(c) $\lambda = 1$ we have a constant population.

**Observation 4.2.1** *A Malthusian process with constant growth ratio $\lambda$ and initial condition $N_0$ is given by*

$$N_n = N_0 \cdot \lambda^n. \tag{4.44}$$

### 4.2.2 The logistic equation

In nature an exponential growth, as described above, cannot go on indefinitely because several limiting factors of the enviroment, for instance lack of food, oxygen, space, light etc. or simply the adverse effects of overcrowding, slows down growth sooner or later. The simplest model is the following: To put it very informally, the key idea is that the rate of growth of the population is directly proportional to "room" available in the population's habitat.

There are two ways to describe the situation mathematically. Suppose $C$ is some constant that describes the total saturation point of the habitat. Then

[4] The use of discrete time is sometimes rather artificial, but it may be appropriate if the the population is censused in intervals, or in generations.

for a population of size $N_n$, we can say that the amount of elbow room is the difference between this capacity and the population size, namely $C - N_n$. When the growth rate is proportional to the amount of elbow room we have

$$\text{growth rate for a period } n = r(C - N_n) + 1,$$

where $r$ is a constant. Altogether we have

$$\begin{aligned} N_{n+1} &= \text{ population at period } n+1 \\ &= (\text{ population at period } n) \cdot (\text{growth rate for a period } n) \\ &= N_n(r \cdot (C - N_n) + 1). \end{aligned}$$

Normalizing this equation we get the logistic equation

$$N_{n+1} = cN_n(1 - N_n). \tag{4.45}$$

This equation is sometimes known as the Verhulst equation.[5]

### 4.2.3 Age-structured populations

Fibonacci's rabbits are age-structured, that means their age is very important in determining their vital parameters. In this case the only vital parameter is the birth rate, as none of the rabbits ever die. If we define

(i) $N_{1,n}$ to be the number of one-month-old pairs of rabbits at time $n$; and

(ii) $N_{2,n}$ to be the number of adult pairs of rabbits at time $n$.

Of course,

$$\begin{aligned} N_{1,n+1} &= N_{2,n} && (4.48) \\ N_{2,n+1} &= N_{1,n} + N_{2,n}. && (4.49) \end{aligned}$$

Hence, in matrix notation

$$\begin{pmatrix} N_{1,n+1} \\ N_{2,n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} N_{1,n} \\ N_{2,n} \end{pmatrix}. \tag{4.50}$$

[5] In general, this equation is given in differential form:

$$\frac{dN}{dt} = cN(1 - N), \tag{4.46}$$

with solution

$$N = \frac{N(0)e^{ct}}{1 - N(0) + N(0)e^{ct}}, \tag{4.47}$$

compare [11] or [32].

Generalizing, we may have to consider a (matrix-) recurrence-relation

$$N_{n+1} = LN_n, \tag{4.51}$$

where $N_n$ is a $m$-vector and $L$ is a $m \times m$-matrix known as Leslie matrix, with $N_0$ given.
If $L$ is constant, we are interested in establishing when the process is stationary, that means we consider $L^n$ when $n$ runs to infinity.
Let us look for a solution in the form $N_n = \lambda^n N$. Substituting this in (4.51) and dividing the equation by $\lambda^n$ gives

$$\lambda N = LN \tag{4.52}$$

or, equivalently,

$$(L - \lambda E)N = o. \tag{4.53}$$

Of course we are interested in non-trivial solution for this equation; we look for an Eigenvalue of $L$. For this to happen the matrix $L - \lambda E$ must be singular, which implies

$$\det(L - \lambda E) = 0. \tag{4.54}$$

This is a polynomial of degree $m$ and has $m$ roots.

**Theorem 4.2.2** *Let the recurrence equation (4.51) with a Leslie matrix $L$ be given. Assuming that all eigenvalues $\lambda_1, \ldots, \lambda_m$ of $L$ are distinct, then a general solution of the recurrence relation is*

$$N_n = \sum_{i=1}^{m} \alpha_i \lambda_i^n c_i, \tag{4.55}$$

*where $c_i$ is the eigenvector corresponding to $\lambda_i$. The numbers $\alpha_i$ are constants that are determined by the initial conditions.*

For further information compare [11].

## 4.3 NUMERICAL RECURRENCES

### 4.3.1 The arithmetic-geometric mean

Let $x_1, \ldots, x_n$ be nonnegative real numbers. The geometric mean for these numbers is defined by

$$G(x_1, \ldots, x_n) = \sqrt[n]{x_1 \cdots x_n}, \tag{4.56}$$

and the arithmetic mean by

$$A(x_1,\ldots,x_n)=\frac{x_1+\ldots+x_n}{n}. \tag{4.57}$$

Many different and ingenious proofs of the very important inequality $G \leq A$ have been devised. The simplest way is the following.

**Theorem 4.3.1** *Let $x_1,\ldots,x_n$ be nonnegative real numbers. Then*

$$G(x_1,\ldots,x_n)\leq A(x_1,\ldots,x_n), \tag{4.58}$$

*where equality holds if and only if $x_1=\ldots=x_n$.*

*Proof.* Let $s=\sum_{i=1}^n x_i$. Consider the function

$$f(x_1,\ldots,x_n)=\prod_{i=1}^n x_i. \tag{4.59}$$

Since the set $\{(x_1,\ldots,x_n): x_i\geq 0, \sum_{i=1}^n = s\}$ is compact, the quantity Max$f$ exists.
We may assume that all $x_i$ are positive.
Assume that $x_1\neq x_2$, then for $y_1=y_2=\frac{x_1+x_2}{2}$ it holds

$$y_1+y_2+x_3+\ldots+x_n=s, \tag{4.60}$$

and

$$y_1y_2-x_1x_2=\left(\frac{x_1+x_2}{2}\right)^2-x_1x_2=\left(\frac{x_1-x_2}{2}\right)^2>0. \tag{4.61}$$

Consequently, $y_1y_2x_3\cdots x_n > x_1\cdots x_n$. In the same way we can prove that $x_1=x_i$, where $x_i$ is any one of the $x$'s and we may assume that Max$f$ is achieved if $x_i=x_j=x$ for all $i,j$. Then $s=nx$.

$$\begin{aligned}\prod_{i=1}^n x_i &= f(x_1,\ldots,x_n)\\ &\leq f\left(\frac{s}{n},\ldots,\frac{s}{n}\right)\\ &= \left(\frac{s}{n}\right)^n\\ &= \left(\frac{\sum_{i=1}^n x_i}{n}\right)^n.\end{aligned}$$

This gives the assertion.

□

In particular,

$$\sqrt{ab} \leq \frac{a+b}{2}. \tag{4.62}$$

Consider the following recurrence relation

$$\begin{aligned} a_{n+1} &= \frac{1}{2}a_n + \frac{1}{2}g_n; &(4.63)\\ g_{n+1} &= \sqrt{a_n \cdot g_n} &(4.64) \end{aligned}$$

with $a_1 = a$ and $g_1 = b$.
In view of (4.62) we have $g_2 \leq a_2$, and moreover, $g_n \leq a_n$ for all numbers $n$. Consequently

$$\begin{aligned} a_{n+1} &= \frac{1}{2}a_n + \frac{1}{2}g_n \\ &\leq \frac{1}{2}a_n + \frac{1}{2}a_n \\ &= a_n, \end{aligned}$$

and

$$\begin{aligned} g_{n+1} &= \sqrt{a_n \cdot g_n} \\ &\geq \sqrt{g_n \cdot g_n} \\ &= g_n, \end{aligned}$$

which shows that $\{a_n\}$ is a monotonic decreasing sequence, and $\{g_n\}$ a monotonic increasing sequence, and consequently both are convergent sequences. We have

$$a_n = \frac{g_{n+1}^2}{g_n},$$

which implies

**Theorem 4.3.2** *The sequences defined by the recurrence relations*

$$\begin{aligned} a_{n+1} &= \frac{1}{2}a_n + \frac{1}{2}g_n; &(4.65)\\ g_{n+1} &= \sqrt{a_n \cdot g_n} &(4.66) \end{aligned}$$

*with $a_1 = a$ and $g_1 = b$ converge to the same limit, called the arithmetic-geometric mean.*

As an exercise determine the limit of the sequence $a_n$ for which the general term is the arithmetic mean of its two preceding terms: $a_{n+2} = A(a_n, a_{n+1})$.

### 4.3.2 The square root

Let $a$ be a positive real number whose square root is to be found. Suppose $x$ is an estimate of the square root arrived at by some means. If $x$ is greater than $\sqrt{a}$, then $a/x$ will be smaller than $\sqrt{a}$, and conversely. The product of $x$ and $a/x$ is $a$. These two numbers serve as reciprocal estimates of $\sqrt{a}$. Newton's algorithm iteratively replace the estimate $x$ by the means of the reciprocal estimates:

$$\begin{aligned} x_1 &= a && (4.67) \\ x_{n+1} &= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), && (4.68) \end{aligned}$$

until the difference of the reciprocal estimates has been made as small as desired. In other terms

$$\lim_{n\to\infty} x_n = \sqrt{a}. \tag{4.69}$$

# 5

# COMBINATORIAL PROBABILITY

The theory of probability is one of the most important areas of mathematics as regards applications. In this book our only goal is to illustrate the importance of combinatorial results by explaining several key facts of the theory of probability.[1]

Historically, counting problems have been closely associated with probability. Indeed, any problem of the kind ”How many objects are there which ...” has the closely related form ”What fraction of all objects ...”, which in turn can be posed as ”What is the probability that a randomly chosen object ...?” when expressed in terms of the theory of probability. In this sense Laplace defined probability as follows

$$\text{Probability } = \frac{\text{number of favorable cases}}{\text{total number of cases}}. \tag{5.1}$$

Probability will be a measure of how likely it is that some event will occur, given as a number between 0 (=impossible) and 1 (=certain).
In the present book we only deal with probability problems where Laplace's definition of probability applies.

## 5.1 EVENTS AND PROBABILITIES

Let $S = \{s_1, \ldots, s_n\}$ be the set of possible outcomes of an experiment. To get a probability space we assume that each outcome $s_i \in S$ has a probability $p(s_i)$

[1] And its applications in biomathematics.

such that

$$\begin{aligned} \sum_{i=1}^{n} p(s_i) &= 1 \quad \text{and} & (5.2)\\ p(s_i) &\geq 0, & (5.3) \end{aligned}$$

for all $i = 1, \ldots, n$. Consequently

$$p(s_i) \leq 1, \tag{5.4}$$

for all $i = 1, \ldots, n$.

A subset $E$ of $S$ is called an event. The probability of an event $E \subseteq S$ is defined as the sum of probabilities of outcomes in $E$, and is denoted by $p(E)$:

$$p(E) = \sum_{s \in E} p(s). \tag{5.5}$$

A probability space in which every outcome has the same probability is called a uniform probability space. In this case the probability of an event $E$ is

$$p(E) = \frac{|E|}{|S|}. \tag{5.6}$$

## 5.2 THE ALGEBRA OF PROBABILITIES

The following theorem follows directly from the Laplace definition, and gives the characteristic properties for probability.

**Theorem 5.2.1** *The probability function p defined on the class of all events in a finite probability space has the following properties:*

*(a)* *For every event $E$, $0 \leq p(E) \leq 1$.*

*(b)* *$p(S) = 1$.*

*(c)* *If events $E$ and $F$ are mutually exclusive, then $p(E \cup F) = p(E) + p(F)$.*

The concept of the algebra of sets enters into the calculation of probabilities when the probabilities of certein events are known and the probability of others

are required.
For example, from the knowledge of $p(E)$, $p(F)$ and $p(E\cap F)$ we may compute the probability of $p(E\cup F)$. In view of 1.42 we have derive the following law.

**Theorem 5.2.2** *Let $E$ and $F$ be events, then*

$$p(E\cup F) = p(E) + p(F) - p(E\cap F). \tag{5.7}$$

As an example toss a coin three times and observe the sequence of heads (H) and tails (T) that appears. The probability space consists of 8 elements. Let $E$ be the event that two or more heads appear consecutively, and $F$ that all the tosses are the same, that means:

$$\begin{aligned} E &= \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{THH}\}, \\ F &= \{\mathrm{HHH}, \mathrm{TTT}\}. \end{aligned}$$

Then

$$E\cap F = \{\mathrm{HHH}\}$$

is the event in which only heads appear. The probilities are

$$p(E) = \frac{3}{8},\ p(F) = \frac{1}{4},\ \text{and } p(E\cap F) = \frac{1}{8}.$$

This last theorem can be generalized to the following equation, which follows immediately from 1.5.3.

**Theorem 5.2.3** *Let $E_1, \ldots, E_n$ be a (finite) collection of events. Then*

$$\begin{aligned} p(\bigcup_{i=1}^{n} E_i) &= \sum_{i=1}^{n} p(E_i) - \sum_{1\le i<j\le n} p(E_i\cap E_j) \\ &\quad + \sum_{1\le i<j<k\le n} p(E_i\cap E_j\cap E_k) \mp \ldots - (-1)^n p(\bigcap_{i=1}^{n} E_i). \end{aligned} \tag{5.8}$$

The following facts for events $E$ and $F$ are easily to see:

(a) $p(E^c) = 1 - p(E)$.

(b) $p(\emptyset) = 0$.

(c) $p(E \setminus F) = p(E) - p(E \cap F)$.

(d) If $E \subseteq F$, then $p(E) \leq p(F)$.

Although it is very simple, the following result from our considerations, is tremendously useful.

**Corollary 5.2.4** *For any finite or countable infinite collection $E_{\alpha \in A}$ of events*

$$p(\bigcup_{\alpha \in A} E_\alpha) \leq \sum_{\alpha \in A} p(E_\alpha). \tag{5.9}$$

## 5.3 CONDITIONAL PROBABILITY AND INDEPENDENT EVENTS

Baye's theorem says that if an event $E$ is actually observed, then the probability of an hypothesis $H$ must be multiplied by the following ratio:

$$\frac{\text{probability of observing } E \text{ if } H \text{ is true}}{\text{probability of observing } E}. \tag{5.10}$$

In other words, the conditional probability of an hypothesis $H$ given an event $E$ is equal to the ratio of the unconditional probability of $H$ multiplied by the conditional probability of $E$ if $H$ is true to the unconditional probability of $E$ alone. More formally, suppose $E$ is an event in a probability space $(S, p)$ with $p(E) > 0$. The probability that an event $F$ occurs once $E$ has occured, called the conditional probability of $F$ given $E$, written $p(F|E)$, is defined as follows

$$p(F|E) = \frac{p(F \cap E)}{p(E)}. \tag{5.11}$$

Roughly spoken, $p(F|E)$ measures the relative probability of $F$ with respect to the reduced space $E$. According to Laplace this is defined by

$$\text{Conditional Probability } = \frac{\text{number of cases in } F \cap E}{\text{total number of cases in } E}. \tag{5.12}$$

Multipying both sides of (5.11) by $p(E)$ gives us the following multiplication theorem for the conditional probability.

**Theorem 5.3.1** *Let $E, F$ be events in a probability space $(S, p)$.*

$$p(F \cap E) = p(F|E) \cdot p(E). \tag{5.13}$$

The multiplication theorem gives us a formula for the probability that events $E$ and $F$ both occur. It can easily be extended to more than two events $E_1, \ldots, E_n$; that is

$$p(E_1 \cap E_2 \cap \ldots \cap E_n) = p(E_1) \cdot p(E_2|E_1) \cdots p(E_n|E_1 \cap E_2 \cap \ldots \cap E_{n-1}). \tag{5.14}$$

Events $E$ and $F$ in a probability space $(S, p)$ are said to be independent if the occurence of one of them does not influence the occurence of the other. More exactly, $F$ is independent of $E$ if $p(F)$ is the same as $p(F|E)$. Substituting this in 5.3.1 yields

**Theorem 5.3.2** *Two events $E$ and $F$ in a probability space $(S, p)$ are independent if*

$$p(F \cap E) = p(F) \cdot p(E). \tag{5.15}$$

Recall our probability space from above; a coin is tossed three times. Consider the events:

$$\begin{aligned} E = \{\text{first toss is heads}\} &= \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}\} \\ F = \{\text{second toss is heads}\} &= \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{THH}, \mathrm{THT}\} \\ G = \{\text{exactly two heads consecutively}\} &= \{\mathrm{HHT}, \mathrm{HHT}\}. \end{aligned}$$

It is easy to compute that

$$\begin{aligned} p(E) \cdot p(F) &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = p(E \cap F), && \text{hence } E \text{ and } F \text{ are independent} \\ p(E) \cdot p(G) &= \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} = p(E \cap G), && \text{hence } E \text{ and } G \text{ are independent} \\ p(F) \cdot p(G) &= \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \neq p(F \cap G), && \text{hence } F \text{ and } G \text{ are dependent} \end{aligned}$$

## 5.4 THE BIRTHDAY PARADOX

Let $p_n$ be the probability that any two of $n$ persons picked at random have the same birthday. It is easier to first compute the probability that no two of the

$n$ persons have the same birthday. Here we think of a list of the people, and ask: ”In how many ways is it possible for the birthdays of each in turn to be different from those above them on the list?” Then

$$1 - p_n = \frac{365 \cdot 364 \cdot 363 \cdots (365 - n + 1)}{365^n}. \tag{5.16}$$

Consequently,

| $n$ | $p_n$ |
|---|---|
| 2 | 0.0027 |
| 3 | 0.0082 |
| ⋮ | ⋮ |
| 20 | 0.4114 |
| ⋮ | ⋮ |
| 41 | 0.9032 |
| ⋮ | ⋮ |
| 57 | 0.9901 |
| ⋮ | ⋮ |
| 80 | 0.9999 |
| ⋮ | ⋮ |
| 365 | 1 |

The strange behavior of $p_n$ is often called the birthday paradox.[2]

Another way of expressing (5.16) is

$$\left(1 - \frac{1}{365}\right) \cdot \left(1 - \frac{2}{365}\right) \cdot \left(1 - \frac{3}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right). \tag{5.17}$$

More generally, if there are $m$ people and $n$ possible birthdays then the probability that all $m$ have different birthdays is

$$\left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \left(1 - \frac{3}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right) = \Pi_{j=1}^{m-1} \left(1 - \frac{j}{n}\right). \tag{5.18}$$

If $j$ is small compared to $n$ we have

$$1 - \frac{j}{n} \approx e^{-j/n}. \tag{5.19}$$

[2] This fact is not really a paradox, it seems surprising because we are used to comparing our particular birthday with others and only rarely finding a perfect match.

Using this approximation, assuming that $m$ is small compared to $n$ we find

$$\begin{aligned}\Pi_{j=1}^{m-1}\left(1-\frac{j}{n}\right) &\approx \Pi_{j=1}^{m-1}e^{-j/n}\\ &= e^{-\sum_{j=1}^{m-1}\frac{j}{n}}\\ &= e^{-m(m-1)/2n}\\ &\approx e^{-m^2/2n}.\end{aligned}$$

Consequently, we proved the following theorem.

**Theorem 5.4.1** *If there are $m$ people and $n$ possible birthdays then the probability that all $m$ have different birthdays is*

$$\Pi_{j=1}^{m-1}\left(1-\frac{j}{n}\right) \approx e^{-m^2/2n}. \tag{5.20}$$

This paradox is a classic example of so-called coincidence. In a world where there are great many potential coincidences each with a small probability of happening, someone, somewhere is going to see one. The fact that there are countless numbers of noncoincidences and many people who do not see a significant coincidence in the same period of time is overlooked. Consequently, we tend to underestimate the probabilities of coincidences in certain situations.

## 5.5 RANDOM VARIABLES

When studying a random event, we are often interested in some value associated with the event rather than in the event itself. More exactly a random variable $X$ on a probability space $(S, p)$ is a real-valued function on $S$.

For a random variable $X$ and a real number $a$, the event $[X = a]$ includes all the basic events of the space in which $X$ assumes the value $a$:

$$[X = a] = \{s \in S : X(s) = a\}. \tag{5.21}$$

Consequently,

$$p(X = a) = \sum_{s\in S} p(s). \tag{5.22}$$

### 5.5.1 The expectation of a random variable

The expectation of a random variable is a weighted average of the values it assumes, where each value is weighted by the probability that the variable assumes that value.

$$\mathbf{E}[X] = \sum_i i \cdot p(X = i), \tag{5.23}$$

where we sum over all values in the range of $X$.

**Remark 5.5.1** *For any collection of random variables $X_1, \ldots, X_n$ and constants $c_1, \ldots, c_n$,*

$$\mathbf{E}\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} c_i \mathbf{E}[X_i]. \tag{5.24}$$

Markov's inequality, formulated in the next theorem, is a fundamental fact for random variables.

**Theorem 5.5.2** *Let $X$ be a random variable that assumes only nonnegative values. Then for all $a > 0$*

$$p(X \geq a) \leq \frac{\mathbf{E}[X]}{a}. \tag{5.25}$$

*Proof.* Let

$$Y = \begin{cases} 1 & : \quad X \geq a \\ 0 & : \quad \text{otherwise} \end{cases}$$

Since, $X \geq 0$

$$Y \leq \frac{X}{a}. \tag{5.26}$$

Because $Y$ is a 0/1 random variable we have

$$\mathbf{E}[Y] = p(Y = 1) = p(X \geq a).$$

Taking expectations in (5.26) and using the linearity 5.5.1 thus yields

$$p(X \geq a) = \mathbf{E}[Y] \leq \mathbf{E}\left[\frac{X}{a}\right] = \frac{\mathbf{E}[X]}{a}.$$

□

### 5.5.2 The variance of a random variable

The variance of a random variable $X$ offer a measure of how far the random variable is likely to be from its expectation.

$$\mathbf{V}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]. \tag{5.27}$$

We give another form for the variance.

**Observation 5.5.3** *The variance of a random variable $X$ equals*

$$\mathbf{V}[X] = \mathbf{E}[X^2] - \mathbf{E}[X]^2. \tag{5.28}$$

*Proof.* Keeping in mind that $\mathbf{E}[X]$ is a constant and 5.5.1, we have

$$\begin{aligned}
\mathbf{V}[X] &= \mathbf{E}[(X - \mathbf{E}[X])^2] \\
&= \mathbf{E}[X^2 - 2X\mathbf{E}[X] + \mathbf{E}[X]^2] \\
&= \mathbf{E}[X^2] - 2\mathbf{E}[X\mathbf{E}[X]] + \mathbf{E}[X]^2 \\
&= \mathbf{E}[X^2] - 2\mathbf{E}[X]\mathbf{E}[X] + \mathbf{E}[X]^2 \\
&= \mathbf{E}[X^2] - \mathbf{E}[X]^2.
\end{aligned}$$

□

Using the expectation and the variance of a random variable, one can derive a strong tail bound, known as Chebyshev's inequality.

**Theorem 5.5.4** *Let $X$ be a random variable. Then for all $a > 0$*

$$p(|X - \mathbf{E}[X]| \geq a) \leq \frac{\mathbf{V}[X]}{a^2}. \tag{5.29}$$

*Proof.* We first observe that

$$p(|X - \mathbf{E}[X]| \geq a) = p((X - \mathbf{E}[X])^2 \geq a^2). \tag{5.30}$$

Since $(X - \mathbf{E}[X])^2$ is a nonnegative random variable, we can apply 5.5.2 to see

$$p((X - \mathbf{E}[X])^2 \geq a^2) \leq \frac{\mathbf{E}[(X - \mathbf{E}[X])^2]}{a^2} = \frac{\mathbf{V}[X]}{a^2}. \tag{5.31}$$

□

### 5.5.3 Bernoulli and Binomial random variables

Suppose that we run an experiment that succeeds with probability $p$ and (consequently) fails with probability $1 - p$.

$$Y = \begin{cases} 1 & : \text{ the experiment succeeds} \\ 0 & : \text{ otherwise} \end{cases}$$

is called a Bernoulli random variable.

**Theorem 5.5.5** *Let $Y$ be a Bernoulli random variable with probability $p$. Its expectation equals*

$$\mathbf{E}[Y] = p. \tag{5.32}$$

*Proof.*

$$\mathbf{E}[Y] = 1 \cdot p + 0 \cdot (1 - p) = p.$$

□

Consider a sequence of $n$ independent experiments, each of which succeeds with probability $p$. If we let $X$ represent the number of successes in the experiments, then $X$ has a binomial distribution, and can be defined by

$$p(X = j) = \binom{n}{j} p^j (1 - p)^{n-j}. \tag{5.33}$$

That means $X$ equals $j$ when there are exactly $j$ successes and $n - j$ failures in $n$ experiments. In view of the binomial theorem we have

$$\sum_{j=0}^{n} p(X = j) = 1, \tag{5.34}$$

As an example consider that a coin is tossed three times; call a heads a succes. This is a binomial distributed experiment $X$ with $n = 3$ and $p = 0.5$. The probability that exactly two heads occur is

$$p(X = 2) = \binom{3}{2} \cdot 0.5^2 \cdot 0.5^1 = 0.365.$$

**Theorem 5.5.6** *Let $X$ be a binomially distributed random variable with parameters $n$ and $p$. Its expectation equals*

$$\mathbf{E}[X] = np. \tag{5.35}$$

*Proof.*

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j} \\
&= \sum_{j=0}^{n} j \frac{n!}{j!(n-j)!} p^j (1-p)^{n-j} \\
&= \sum_{j=1}^{n} \frac{n!}{(j-1)!(n-j)!} p^j (1-p)^{n-j} \\
&= np \sum_{j=1}^{n} \frac{(n-1)!}{(j-1)!((n-1)-(j-1))!} p^{j-1} (1-p)^{(n-1)-(j-1)} \\
&= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!((n-1)-k)!} p^k (1-p)^{(n-1)-k} \\
&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} \\
&= np,
\end{aligned}$$

where 2.2.3 is used to get the last equation.

□

The variance of a binomial random variable can be determined directly by computing $\mathbf{E}[X^2]$ and using 5.5.3 (Exercise).

**Theorem 5.5.7** *Let $X$ be a binomially distributed random variable with parameters $n$ and $p$. Its variance equals*

$$\mathbf{V}[X] = np(1-p). \tag{5.36}$$

### 5.5.4 The law of large numbers

In this section we study an experiment that consits of $n$ independent attempts and has two possible outcomes with the same probability. For instance:

(a) Tossing a coin.

(b) "Walking on the integers"; with each step one position to the left or one position to the right.

In the second example the probability of your position is given by the following table.

| after | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| step 0 | | | | | | | 1 | | | | | | |
| step 1 | | | | | | $\frac{1}{2}$ | | $\frac{1}{2}$ | | | | | |
| step 2 | | | | | $\frac{1}{4}$ | | $\frac{2}{4}$ | | $\frac{1}{4}$ | | | | |
| step 3 | | | | $\frac{1}{8}$ | | $\frac{3}{8}$ | | $\frac{3}{8}$ | | $\frac{1}{8}$ | | | |
| step 4 | | | $\frac{1}{16}$ | | $\frac{4}{16}$ | | $\frac{6}{16}$ | | $\frac{4}{16}$ | | $\frac{1}{16}$ | | |
| step 5 | | $\frac{1}{32}$ | | $\frac{5}{32}$ | | $\frac{10}{32}$ | | $\frac{10}{32}$ | | $\frac{5}{32}$ | | $\frac{1}{32}$ | |
| step 6 | $\frac{1}{64}$ | | $\frac{6}{64}$ | | $\frac{15}{64}$ | | $\frac{20}{64}$ | | $\frac{15}{64}$ | | $\frac{6}{64}$ | | $\frac{1}{64}$ |
| ⋮ | | | | | | | | | | | | | ⋱ |

The law of large numbers says that if we carry out the experiment many times, the frequency by which we get a result will be the same for both possible outcomes:

(a) The number of "heads" will be about the same as the number of "tails".

(b) We are at the first integer.

A more precise formulation is given by the following theorem.

**Theorem 5.5.8** *Let $0 \leq t \leq m$. Then the probability that from $2m$ attempts of an experiment with two outcomes (with the same probability), the number of one outcames is less than $m - t$ or larger than $m + t$ is at most $e^{-t^2/(m+t)}$.*

*Proof.* For simplicity we call the two outcomes head and tail.
Let $E_k$ denote the event that we get exactly $k$ heads. It is clear that the events

$E_k$ are mutually exclusive, and that for every outcome of the experiments, exactly one of the $E_k$ occurs.
The number of outcomes for which $E_k$ occurs is the number of sequences of length $n$ consisting of $k$ heads and, consequently $n-k$ tails. This can be achieved in $\binom{n}{k}$ ways, so that $E_k$ has $\binom{n}{k}$ many elements. Since the total number of outcomes is $2^n$, we get

$$p(E_k) = \frac{\binom{n}{k}}{2^n}. \tag{5.37}$$

Now we calculate the probability that the number of heads is far from the expected $m = n/2$. Assume that it is less than $m-t$ or larger than $m+t$, with $0 < t \leq m$. The probability that this happens is

$$\frac{1}{2^{2m}}\left(\binom{2m}{0} + \ldots + \binom{2m}{m-t-1} + \binom{2m}{m+t+1} + \ldots + \binom{2m}{2m}\right). \tag{5.38}$$

Then 2.2.10 proves the theorem.

□

### 5.5.5 The Poisson random variables

Let $X$ be a binomially distributed random variable with parameter $n$ and $p$. If the number $n$ of experiments is large and the number $k$ of successes is small, then a good approximation is

$$p(X=k) \approx \frac{(np)^k}{k!}e^{-np}, \tag{5.39}$$

assuming that $np$, the expected number of successes, is a constant. This can be seen by the following chain of equations.

$$\begin{aligned}
&\binom{n}{k}p^k(1-p)^{n-k} \\
= \quad & \frac{n(n-1)\cdots(n-k+1)}{k!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} \qquad \text{substituting } p = \frac{\lambda}{n} \\
= \quad & 1\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\cdot\frac{\lambda^k}{k!}\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-k} \\
\to \quad & \frac{\lambda^k}{k!}\cdot e^{-\lambda},
\end{aligned}$$

provided that $n$ increases, $p$ approaches 0, and $\lambda = np =$ const. Consequently, we proved

**Theorem 5.5.9** *(Poisson)*

$$\lim_{n\to\infty, p\to 0, np=\lambda=const} \binom{n}{k} p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} \cdot e^{-\lambda}. \tag{5.40}$$

In view of this theorem we introduce the following notation: A random variable $X$ has the Poisson distribution with parameter $\lambda$ if

$$p(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{5.41}$$

In other terms,

**Observation 5.5.10** *Let $B(k,n,p)$ be binomial distributed with the parameters $n$ and $p$, then*

$$B(k,n,p) \approx P(k,\lambda), \tag{5.42}$$

*where $P(k,\lambda)$ is Poisson distributed with parameter $\lambda = np$.*

This observation is very helpful, because, in general, the quantity $B(k,n,p)$ is hard to compute if $n \gg 1$ and $p \ll 1$.

The Poisson distribution is often used in modelling situations in biology where events occur infrequently. Consider the following example [22]: From many studies, it has become clear that the rate of amino acid substitution varies between organisms and also between protein classes. We are interested in the way how amino acid substitution rates are computed.
Let $w$ and $w'$ be two (homologous) polypeptides of the same length $n$. $n_d$ denotes the number of differences between homologous acid sites; the probability $p$ of an amino acid substituting occuring at a given site of either $w$ or $w'$ can be estimated by

$$p \approx \frac{n_d}{n}. \tag{5.43}$$

A second approximation of $p$ can be derived by assuming that the substitution of amino acids at a given site is a Poisson process. Let $X$ be a random variable counting the number of mutations over time $t$ at fixed site for an polypeptide having substitution rate $\lambda$ per site (and per year). Then

$$p(X = k) = \frac{(\lambda \cdot t)^k}{k!} e^{-\lambda \cdot t}. \tag{5.44}$$

Thus the probability that no substitution occurs at a given site in $w$ is

$$p(X=0)=e^{-\lambda\cdot t}. \tag{5.45}$$

Hence the probability that no substitution occurs at a given site in $w$ and $w'$ is

$$q=e^{-2\cdot\lambda\cdot t}. \tag{5.46}$$

Since $d=2\cdot\lambda\cdot t$ is the total number of substitutions occuring at a fixed site, we get

$$d=2\cdot\lambda\cdot t=-\ln q. \tag{5.47}$$

Together with (5.43) we find the following approximation

$$d\approx-\ln\left(1-\frac{n_d}{n}\right) \tag{5.48}$$

for the protein substitution rate.

**Theorem 5.5.11** *Let $X$ be a Poisson distribution with parameter $\lambda$. Its expectation equals*

$$\mathbf{E}[X]=\lambda. \tag{5.49}$$

*Proof.* Recall that the $k$th term in the Taylor expansion of $e^x$ is $x^k/k!$, so that

$$\begin{aligned}\mathbf{E}[X] &= \sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}\\ &= \lambda e^{-\lambda}\frac{d}{d\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\\ &= \lambda e^{-\lambda}\frac{d}{d\lambda}e^{\lambda}\\ &= \lambda e^{-\lambda}e^{\lambda}\\ &= \lambda.\end{aligned}$$

□

The variance can be determined by considering the function $\mathbf{E}[e^{tX}]$ and its derivations (Exercise).

**Theorem 5.5.12** *Let $X$ be a Poisson distribution with parameter $\lambda$. Its variance equals*

$$\mathbf{V}[X]=\lambda. \tag{5.50}$$

## 5.6 RANDOM GENETIC DRIFT

Natural selection is not the only factor that can cause changes in allele frequencies, it can also occur by chance, in which case the changes are not directional but random. An important factor in producing random fluctuations in allele frequencies is the random sampling of gametes in the process of reproduction. Let us consider a population in which

(a) all individuals have the same fitness;

(b) selection does not operate;

(c) the generations are nonoverlapping; and

(d) the population size does not change from generation to generation.

The population under consideration is diploid and consists of $n$ individuals, so that at any given locus the population contains $2n$ genes. Let $p$ be the frequency of allele $A$ in the population and $1-p$ be the frequency of the allele $B$, respectively. A random variable $X$ is given when $2n$ gametes are sampled from the gamete pool, the probability that the sample contains exactly $i$ alleles of type $A$ is given by

$$p(X=i)=\frac{(2n)!}{i!(2n-i)!}p^i(1-p)^{2n-i}. \tag{5.51}$$

The process of change in allele frequency due solely to chance effects is called random genetic drift. Let us follow the dynamics of chance of the frequencies

$$p_0, p_1, \ldots, p_t, \ldots, \tag{5.52}$$

of the allele $A$ in succeeding generations, where the subscripts denote the generation number.
On average $p_1$ will be equal to $p_0$, and furthermore $p_{i+1} \approx p_i$, for all $i$. In reality, sampling occurs only once in each generation, namely

(i) $p_1$ is usually different from $p_0$;

(ii) the frequency $p_2$ will no longer depend on $p_0$, but only on $p_1$;

(iii) the frequency $p_3$ will depend on neither $p_0$ nor $p_1$, but only on $p_2$; and so on.

Thus the most important property of the genetic drift is its cumulative behavior. In mathematical terms, we are interested in the expectation and the variance of the frequency of allele $A$ in generation $t$. Without proof we give these quantities by

$$\begin{aligned} \mathbf{E}_t &= p_0 && (5.53) \\ \mathbf{V}_t &\approx p_0(1-p_0)(1-e^{-t/2n}). && (5.54) \end{aligned}$$

Note that although the expectation mean does not change with time, the variance increases. In other terms, the chance in allele frequencies is not systematic in its direction.

## 5.7 WATSON'S PARADOX

Clearly, it is of great interest to understand the evolutionary past of mankind, to specify the location of the human branch of the tree of life. This is one of the biggest questions in evolutionary biology.

Darwin claimed that the African apes are mans closest relatives, and suggested that evolutionary origins of man were to be found in Africa. In other words, the commonly held view was that humans were phylogenetically distinct from the great apes (chimpanzees, gorillas and orang-utans), being placed in different taxonomic families, and that this split occurred at least 15 Mya. These conclusions were based on fossils.

Genetic studies of human prehistory started 100 years ago considering blood groups. By 1964, knowing much more about blood groups and their worldwide distributions, Cavalli-Sforza and Edwards constructed the first family tree of human species.

In 1967, Sarich and Wilson [72] measured the extent of immunological cross-reaction in the protein serum albumin between various primates. The results were striking: humans, chimpanzees and gorillas were genetically equidistant and clearly distinct from the orang-utan.[3]

There are two different models:

- The multiregional model posits the evolution of *Homo sapiens* from a convergence of various distinct hominid lines in different geographic regions.
- The Out of Africa model posits the evolution of a lineage of hominids who left Africa not more than 1 Mya.

[3] The work of Sarich and Wilson was one of the first examples of molecular systematics, that is the use of gene and protein sequences to reconstruct the evolutionary history. It changed the perspective on human origins and opened the "molecules versus morphology" debate.

For more facts about this question compare [6], [7], and [63].
The breakthrough for understanding came with a publication in *Nature* in 1987 [14] by the late Wilson and two of his students, Cann and Stoneking, entitled "Mitochondrial DNA and human evolution". They used mother-only genes, known technically as mitochondrial DNA. Wilson and his colleagues examined the mother-only genes in 134 individuals from around the world. They found remarkable similarities as well as differences in all the samples. The centrepiece of the article was a diagram which bears a superficial resemblance to a tree. It contains a hypothetical common female ancestor of all extant humans, called Eve, or in more scientific terms Mitochondrial Eve (mtEve).
More formally, consider a population of a finite number of individuals who are reproducing independently by the same probability distribution. The probability an individual having $r$ offsprings in the next generation is $p_r$, $r = 0, 1, \ldots, n$, where $n$ being a chosen maximal number of offsprings. Additionally we assume for simplicity that the generations are discrete and each individual has the same life span. The random variable $X_i$ is the number of individuals in the $i$th generation, where $i = 0, 1, 2, \ldots$.
The sequence

$$X_0, X_1, X_2, \ldots \tag{5.55}$$

is called a branching process.
We are interested in the question: What is the probability of a population dying out?

**Theorem 5.7.1** *Let $\kappa = \mathbf{E}[X]$ be the expectation of the number of offspring of an individual. Then for the population to survive the following holds.*

*(a)* *If $\kappa \leq 1$ then the population becomes extinct with probability 1.*

*(b)* *If $\kappa > 1$ then the population becomes extinct with probability $q$, where $q$ is the (unique) root of the polynomial*

$$\Phi(x) - x = \sum_{r=0}^{n} p_r x^r - x. \tag{5.56}$$

For a proof compare [32], [52], [53].

To conclude, it seems most likely that anatomically modern humans evolved in Africa around 200 kya (kilo years ago) and then spread around the world.[4]

[4] As reported in [47], Templeton later obtained several distinct trees, similar to Wilson's tree, and most of them support a non-African hypothesis. But the "Out of Africa" hypothesis is also supported by several other observations.

For a informative discussion on this subject see also [5], [13], [17], [18], [61], and [90].

## 5.8 THE THEORY OF INFORMATION

In everyday life, we use the word "information" in many different ways. One of the most common use is as a measure of novelty or surprise. In a broad sense, coding theory is concerned with the transfer of information- one with reliability, the other with security.
Compare [74] for a common description of information and coding theory, and [15] or [91] for its application in molecular biology.

### 5.8.1 Bits

Suppose we want to develop a way to represent the letters of the English alphabet by using words over $\{0, 1\}$. Since there are 26 letters, we should be able to encode these letters in terms of sequences of five bits, given that $2^4 \leq 26 \leq 2^5$. In general we have the following observation.

**Observation 5.8.1** *For an alphabet of $k$ letters we need at least $\lfloor \log k \rfloor$ bits and at most $\lfloor \log k \rfloor + 1$ bits to encode it, where bit means a 0/1-decision.*

We create a mathematical theory for measuring information. Because we do not know in advance what combination of characters might be transmitted, we can only attach a probability to the transmission of any particular character. If the a priori probability of a character $a$ is $p(a)$, then the information content $I(a)$ of $a$ is simply the negative of the logarithm of the probability of that character being sent:

$$I(a) = -\log p(a). \tag{5.57}$$

The base used in calculating the logarithm determines the unit of information. The most common unit used base 2, in which case the information is measured in binary digits (bits).
Now, we will show that (5.57) holds from obvious assumptions.

**Theorem 5.8.2** *Let $A$ be an alphabet with $k \geq 2$ letters and let a probability $p(a)$ of each character be given. Assuming that the information content $I(a)$*

*depends only on the probabilities: $I(a) = f(p(a))$, and the function $f$ satisfies the following two conditions*

*(i) $f$ is continuous; and*

*(ii) $f(p \cdot p') = f(p) + f(p')$ for all $p, p'$.*

*Then we have*

$$f(p) = c \cdot \ln p, \tag{5.58}$$

*where $c$ is a constant.*
*Normalizing $f(1/2) = 1$, implies (5.57).*

*Proof.* Consider the function

$$h = f \circ \exp\,. \tag{5.59}$$

Then $h$ is a continuous function with

$$h(x + y) = f(e^{x+y}) = f(e^x \cdot e^y) = f(e^x) + f(e^y) = h(x) + h(y). \tag{5.60}$$

Under these conditions $h$ must be a linear function: $h(x) = c \cdot x$.[5] This immediately implies the assertion.
The norming procedure gives

$$1 = f\left(\frac{1}{2}\right) = c \cdot \ln \frac{1}{2} = c \cdot (-\ln 2),$$

and

$$f(p) = \frac{-\ln p}{\ln 2} = -\log p.$$

□

[5] For a proof consider successively

(i) $h(0) = 0$.
Since $h(0) = h(0 + 0) = h(0) + h(0)$.

(ii) $h(-x) = -h(x)$.
Since $0 = h(x - x) = h(x) + h(-x)$.

(iii) For positive integers $h(n) = nh(1)$.
Since $h(n) = h(1 + \ldots + 1) = h(1) + \ldots + h(1) = nh(1)$.

(iv) For positive integers $h(1/n) = (1/n)h(1)$.
Since $h(1) = h(n/n) = nh(1/n)$.

(v) $h(n/m) = (n/m)h(1)$, for integers $n$ and $m$.
By (iii) and (iv).

(vi) $h(r) = rh(1)$ for real numbers $r$.
By using that $h$ is continuous.

A noncontinuous function $h$ which satisfies $h(x + y) = h(x) + h(y)$ is very strange, see [78].

### 5.8.2 Entropy

Consider the following example: suppose that we have two biased coins. One comes up heads with probability 3/4, and the other comes up heads with probability 7/8. Which coin produces more randomness per flip?

A system $S = (A, p)$ with an alphabet $A$ and a probability $p$ is called an information source. The average information contained in a source is called its entropy. This is a measure of the uncertainty in a system at a given moment because the more information there is in a system, the greater the uncertainty is in specifying exactly what state the system is in. Boltzmann is today credited with having the notion of entropy as a measure of the disorder present in a collection of objects.

Consider a word $w = a_1 \dots a_n \in A^\star$. What is the information within $w$?
Suppose we receive a character $a$ from $A$, with $k = |A|$, where characters are sent with uniform distribution. In view of (5.57) we have

$$I(a) = -\log \frac{1}{k} = \log k. \tag{5.61}$$

Now assume that $p_i = p(a_i)$, $i = 1, \dots, k$ are the probabilities of outputting characters $a_i$ in a message, where

$$\begin{aligned} \sum_{i=1}^{k} p_i &= 1 \quad \text{and} \\ p_i &> 0. \end{aligned}$$

(We omit the case $p_i = 0$, since here $a_i$ does not really occur.)
Let $n_i = np_i$ be the expected number of occurences of $a_i$ in the random message $w$. Then $w$ belongs with high probability to a set of size

$$N_n = \frac{n!}{n_1! \cdots n_k!}. \tag{5.62}$$

The average information should then equal

$$I = \frac{\log N_n}{n}. \tag{5.63}$$

Applying Stirling's formula yields

$$\ln N_n \quad \approx \quad n \ln n - \sum_{i=1}^{k} np_i \ln(np_i)$$

$$\begin{aligned} &= n\ln n - \ln n \sum_{i=1}^{k} np_i - \sum_{i=1}^{k} np_i \ln p_i \\ &= -n \sum_{i=1}^{k} p_i \ln p_i. \end{aligned}$$

Since ln and log are related by a constant, it follows that

$$\log N_n \approx -n \sum_{i=1}^{k} p_i \log p_i, \tag{5.64}$$

consequently

$$I = \frac{\log N_n}{n} = -\sum_{i=1}^{k} p_i \log p_i, \tag{5.65}$$

which leads to the following definition.

**Theorem 5.8.3** *(Shannon's formula) The entropy $H(S) = H(p_1, \ldots, p_k)$ of an information source $S = (A, p)$ is given by*

$$H(p_1, \ldots, p_k) = -\sum_{i=1}^{k} p_i \log p_i. \tag{5.66}$$

About the behavior of the entropy we have the following fact.

**Theorem 5.8.4** *The entropy-function $H(S) = H(p_1, \ldots, p_k)$ of an information source $S = (A, p)$ has its maximum if and only if*

$$p_1 = \ldots = p_k. \tag{5.67}$$

*Proof.* We consider the function

$$H(S) = H(p_1, \ldots, p_k) = -\sum_{i=1}^{k} p_i \log p_i = -\frac{1}{\ln 2} \sum_{i=1}^{k} p_i \ln p_i \tag{5.68}$$

with subject to $\sum_{i=1}^{k} p_i = 1$, with the help of Lagrange's multiplier. That means we consider the function

$$F(p_1, \ldots, p_k, \lambda) = -\frac{1}{\ln 2} \sum_{i=1}^{k} p_i \ln p_i - \lambda \left( \sum_{i=1}^{k} p_i - 1 \right). \tag{5.69}$$

Then setting the partial derivates to 0, we obtain

$$\begin{aligned}\frac{\partial F}{\partial p_i} &= -\frac{\ln p_i}{\ln 2} - \frac{1}{\ln 2} - \lambda = 0, \text{ and}\\ \frac{\partial F}{\partial \lambda} &= \sum_{i=1}^{k} p_i - 1 = 0.\end{aligned}$$

These equalities give the assertion.

□

### 5.8.3 Codes

Let $A$ and $B$ be alphabets. Then an injective mapping

$$c : A \to \bigcup_{i=1}^{k} B^i \tag{5.70}$$

is called a coding of $A$ by (words of) $B$.
The image of $c$ is called a code. It can be extended to a map

$$c^\star : A^\star \to B^\star, \tag{5.71}$$

which is also called coding, by setting

$$c^\star(a_1 \ldots a_n) = c(a_1) \ldots c(a_n). \tag{5.72}$$

Although $c$ is injective, this must not be true for $c^\star$: Consider $A = \{x, y, z\}$ and $B = \{0, 1\}$ with $c(x) = 0, c(y) = 1$ and $c(z) = 01$. Here is $c^\star(xy) = 01 = c^\star(z)$. Consequently, the injectivity of $c^\star$ must be forced.

If $u, v \in B^\star$ and $w = uv$, then the word $u$ is called a prefix and the word $v$ is called a suffix of $w$. The empty word is prefix and suffix of each word. Any word is prefix and suffix of itself. We use these notations for codes which are uniquely decodable. The main examples are the so-called prefix codes. A collection $c(A)$ of words is called a prefix code if no word in $c(A)$ is the prefix of any other word in $c(A)$. For each prefix code $c^\star$ is injective.

A block code is a code having all its words of the same length; this number of letters is called the length of a code. Of course, a block code is a prefix code.

One example of a block code is the ASCII (American standard code for information interchange), by which computers represent alphanumeric characters. ASCII provides 128 code words over $\mathbb{B} = \{0, 1\}$ for 128 characters. Thus, each such codeword contains seven bits of information. An eighth bit is used for error-detecting.
The famous genetic code hardwired into every cell in your body is a good example for another type of a block code. Because there are four possible nucleic acids: adenine (a), cytosine (c), guanine (g), and uracil (u), that can appear at each location in a code word. 20 amino acids: alanine, arginine, . . ., valine, are coded. Hence, each code word must be of length 3.

| | u | c | a | g | |
|---|---|---|---|---|---|
| u | phenylalanine | serine | tyrosine | cysteine | u |
| | phenylalanine | serine | tyrosine | cysteine | c |
| | leucine | serine | *punctuation* | *punctuation* | a |
| | leucine | serine | *punctuation* | tryptophan | g |
| c | leucine | proline | histidine | arginine | u |
| | leucine | proline | histidine | arginine | c |
| | leucine | proline | glutamine | arginine | a |
| | leucine | proline | glutamine | arginine | g |
| a | isoleucine | threonine | asparagine | serine | u |
| | isoleucine | threonine | asparagine | serine | c |
| | isoleucine | threonine | lysine | arginine | a |
| | methionine | threonine | lysine | arginine | g |
| g | valine | alanine | aspartic acid | glycine | u |
| | valine | alanine | aspartic acid | glycine | c |
| | valine | alanine | glutamic acid | glycine | a |
| | valine | alanine | glutamic acid | glycine | g |

### 5.8.4 Huffman codes

Unique decodability means that there can be only a single interpretation for each code. Roughly speaking, if a code is uniquely decipherable, it cannot have very many short code words.

Let $c : A \to \{0, 1\}^\star$ be a coding.

$$N(c) = \max\{|c(a)| : a \in A\} \tag{5.73}$$

is called the maximum code word length. If there is a given probability $p : A \to \mathbb{R}$, the quantity

$$\overline{N}(c) = \sum_{a \in A} p(a) \cdot |c(a)| \tag{5.74}$$

is called the average code word length. Of course, $\overline{N}(c) \leq N(c)$.

**Theorem 5.8.5** *(Shannon's noiseless coding theorem) Let $S = (A, p)$ be an information source. Then*

*(a) It holds*

$$H(S) \leq \overline{N}(c), \tag{5.75}$$

*for any prefix code $c$.*

*(b) There is a prefix code $c$ such that*

$$\overline{N}(c) < H(S) + 1. \tag{5.76}$$

From this we see that there is a code always calls for less that one bit per symbol more than the entropy.
Furthermore this leads to the obvious question: What is the most efficient coding scheme, the one with the smallest average code-word length, for a given information source. This question was answered by Huffman, who discovered an ingenious scheme, called a Huffman encoding, for creating optimal codes. The essential idea in Huffman's procedure is to systematically assign the shortest code words to the symbols that occur most frequently. More exactly,

1. Find the two smallest probabilities $p(a)$ and $p(b)$;
   Draw a line from each of them to their sum $a|b$ with probability

   $$p(a|b) = p(a) + p(b), \tag{5.77}$$

   which is the probability of either the symbol $a$ or the symbol $b$ being transmitted;

2. Create the new source

   $$S' = (A \cup \{a|b\} \setminus \{a, b\}, p'), \tag{5.78}$$

   where $p'$ is new computed by (5.77);
   Repeat the steps until $|A| = 1$;
   We find a collection of paths from each symbol to a common point;

3. Put a 1 on each upper path and a 0 on each lower path;
4. The Huffman code for a given symbol is then the sequence of bits encountered in tracing back from the the common point.

Compare [74].

Nature uses a similar approach, namely using ”supersymbols”, for the genetic code. In the genetic code each codeword has a length of 6 bits, but this is not necessary, when we additionally use the binary alphabet $A' = \{r, y\}$ in which $r$ codes for a purine ($a$ or $g$), $y$ codes for a pyrimidine ($c$ or $u$), each of 1 bit, and $-$ codes for any one of 0 bit.[6]

| | u | c | a | g |
|---|---|---|---|---|
| u | y: phenylalanine<br>r: leucine | -: serine | y: tyrosine<br>r: *punctuation* | y: cysteine<br>a: *punctuation*<br>g: tryptophan |
| c | -: leucine | -: proline | y: histidine<br>r: glutamine | -: arginine |
| a | y: isoleucine<br>a: isoleucine<br>g: methionine | -: threonine | y: asparagine<br>r: lysine | y: serine<br>r: arginine |
| g | -: valine | -: alanine | y: aspartic acid<br>r: glutamic acid | -: glycine |

Glancing at this structure, it is clear that the genetic code is fault-tolerant, in the sense that transcription errors in the third codon position frequently do not influence the amino acid expressed. This is called the wobble-hypothesis.

Consider the amino acid composition[7]:

[6] $-$ is a ”dummy” letter.

[7] given by the Swiss-Prot protein sequence data bank *www.expasy.ch*

| amino acid | scale values (in %) | coded (in bit) |
|---|---|---|
| alanine | 7.8 | 4 |
| cysteine | 1.57 | 5 |
| aspartic acid | 5.3 | 5 |
| glutamatic acid | 6.59 | 5 |
| phenylalanine | 4.02 | 5 |
| glycine | 6.93 | 4 |
| histidine | 2.27 | 5 |
| isoleucine | 5.91 | 5.5 |
| lysine | 5.93 | 5 |
| leucine | 9.62 | 4.5 |
| methionine | 2.37 | 6 |
| asparagine | 4.22 | 5 |
| proline | 4.85 | 4 |
| glutamine | 3.93 | 5 |
| arginine | 5.29 | 4.5 |
| serine | 6.89 | 4.5 |
| threonine | 5.46 | 4 |
| valine | 6.69 | 4 |
| tryptophan | 1.16 | 6 |
| tyrosine | 3.09 | 5 |

The entropy is calculated in 4.174 and the average code word length in 4.642.

## 5.9 THE ORIGIN OF LIFE AND EIGEN'S PARADOX

Consider the origins of life. Suppose that a polynucleotide molecule is of length $m$. We make the following assumptions:

(i) The average number of copies of itself it produces during its lifetime is $s$.

(ii) Some of the copies may not be exact. For simplicity we assume that all copies are produced at the end of the lifetime, and that every nucleotide in the sequence is copied correctly with the same probability $p$, where $0 < p < 1$.

Each copy is a random variable with probability $p^m$. Copies can be seen as experiments of which there are $s$; more exactly as a binomial distribution with

$$\begin{aligned} & p(\text{exactly } k \text{ correct copies of the polynucleotide}) \\ = & \binom{s}{k}(p^m)^k(1-p^m)^{s-k} \\ = & \binom{s}{k}(p^{mk})(1-p^m)^{s-k}. \end{aligned}$$

In view of 5.5.6 the expectation is

$$sp^m. \tag{5.79}$$

This means that the polynucleotide molecule will have $sp^m$ correct copies by the time it disintegrates.
In other terms, the "genotype" of the polynucleotide may survive only if $sp^m > 1$, or conversely, the polynucleotide dies with probability 1 if $sp^m \le 1$. Equivalently,

$$\log(sp^m) \le \log 1 = 0,$$

rewritten by

$$m \ge -\frac{\log s}{\log p}, \tag{5.80}$$

paying attention to the fact that $\log p < 0$.
In view of $\log p \le p - 1$ we have the following result.

**Observation 5.9.1** *Under the assumptions listed above a polynucleotide dies with certainity if*

$$m \ge \frac{\log s}{1-p}. \tag{5.81}$$

Experiments suggest that RNA replication without any enzymes has approximately an error probability of $1 - p = 0.05$. If we assume that $s = 2$ or $s = 3$, then 5.9.1 implies a sequence of the length of at most 20. This is far too short for protein synthesis. With the help of enzymes the probability of error decreases considerably.

> If a RNA molecules are not sufficiently long, enzymes cannot be synthetized and without enzymes RNA cannot reach the length necessary for enzymes synthetization.

This antimony is often called the error catastrophe or the information crisis. A way out of this trap was suggested by Eigen: Called the hypercycle, it is a mathematically well-founded theory of that which is not unimaginable from the point of view of "natural history". Extremely roughly spoken, this is a loop formed by nucleotides and catalysts which realized a coevolution that leads to life. For a complete discussion compare [32].

## 5.10 SEQUENCE SPACES

### 5.10.1 The Hamming distance

For a word $v \in \{0,1\}^n$ we define the Hamming weight $\mathrm{wt}(v)$ as the number of times the digit "1" occurs in $v$. Clearly, $\mathrm{wt}(v) \leq n$. Moreover

$$\mathrm{wt}(v+w) \leq \mathrm{wt}(v) + \mathrm{wt}(w). \tag{5.82}$$

which is easily to show.

Let $v$ and $w$ be words of length $n$. We define the Hamming distance by

$$\rho_H(v,w) = \mathrm{wt}(v+w) = \mathrm{wt}(v-w). \tag{5.83}$$

Conversely,

$$\mathrm{wt}(v) = \rho_H(v,o), \tag{5.84}$$

where $o = 0^n$.
The Hamming distance between $v$ and $w$ is the number of positions in which $v$ and $w$ disagree. It can be directly generalized to words in $A^n$, for an alphabet $A$:

$$\rho_H((a_1,\ldots,a_n),(b_1,\ldots,b_n)) = |\{i : a_i \neq b_i \text{ for } i = 1,\ldots,n\}|, \tag{5.85}$$

for $a_i, b_i \in A$.

**Theorem 5.10.1** $(A^n, \rho_H)$ *is a metric space.*

### 5.10.2 The number of words in sequence spaces

We count the words in $\mathbb{B}^4 = \{0,1\}^4$ by its Hamming weights.

| weight | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| | 0000 | 0001<br>0010<br>0100<br>1000 | 0011<br>0101<br>0110<br>1011<br>1101<br>1110 | 0111<br>1011<br>1101<br>1110 | 1111 |
| number | 1 | 4 | 6 | 4 | 1 |

$\binom{n}{k}$ is just the number of ways that an unordered collection of $k$ elements can be chosen from a set of $n$ elements. Thus $\binom{n}{r}$ is the number of words in $(I\!\!B^n, \rho_H)$ with weight $r$, $0 \leq r \leq n$. That means

**Lemma 5.10.2** *Let $v$ be a word in $(\{0,1\}^n, \rho_H)$ and let $r$ be an integer with $0 \leq r \leq n$. Then the number of words with a distance of at most $r$ from $v$ is precisely*

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{r}. \tag{5.86}$$

For a nonnegative integer $r$ we defined

$$B_r(v) = \{x \in \{0,1\}^n : \rho_H(x, v) \leq r\} \tag{5.87}$$

as the ball with center $v \in X$ and radius $r$. Then 5.10.2 shows

$$|B_r(v)| = \sum_{k=0}^{r} \binom{n}{k}. \tag{5.88}$$

In particular, $|B_r(v)|$ is independent from the center of the ball.
For $n = 3$ and $v = 110$ we have

$$\begin{aligned} B_0(v) &= \{110\} \\ B_1(v) &= \{110, 010, 100, 111\} \\ B_2(v) &= \{110, 010, 100, 111, 000, 101, 011\} \\ B_3(v) &= I\!\!B^3. \end{aligned}$$

Using 2.2.9 we get

**Theorem 5.10.3** *Let $W$ be a set of words in the hypercube $(\{0,1\}^n, \rho_H)$ with radius $r$, $r \leq n$. Then*

$$|W| \leq \left(\frac{e \cdot n}{r}\right)^r. \tag{5.89}$$

Let $A$ be an alphabet. The metric space $(A^n, \rho_H)$ has a strange property:

(i) On one hand, it is a "big" space, since it contains $|A|^n$ many points;

(ii) On the other hand, it is a "small" space, since its diameter equals $n$:

$$\max\{\rho_H(w, w') : w, w' \in A^n\} = n. \tag{5.90}$$

For some deep consequences of this observation for molecular evolution see Eigen [30].
Similarly to 5.10.2 we have that for a word $v$ in $(A^n, \rho_H)$ and an integer $r$ with $0 \leq r \leq n$, the number of words of distance at most $r$ from $v$ is precisely

$$\sum_{k=0}^{r} \binom{n}{k} \cdot (|A| - 1)^k. \tag{5.91}$$

Consequently,

**Theorem 5.10.4** *Let $W$ be a set of words in $(A^n, \rho_H)$ with radius $r$, $r \leq n$. Then*

$$|W| \leq \left(\frac{e \cdot n \cdot (|A| - 1)}{r}\right)^r. \tag{5.92}$$

## 5.10.3 Measuring the editorial distance

Consider the set $A^\star$ of all words over the alphabet $A$. The edit distance $\rho_L$, between two words of not necessarily equal length is the minimal number of "edit operations" required to change one word into the other, where an edit operation is a deletion, insertion, or substitution of a single letter in either word. This distance is also called Levenshtein distance, since it was introduced by Levenshtein [56] in connection with error correcting codes.
As an example consider the two German words $w$ =APFEL and $w'$ =PFERD, where we have $\rho_L(w, w') = 3$.

In molecular biology the Levenshtein distance is used to measure similarity (homogeneity) of two molecular sequences (say DNA or polypetides).[8]

[8] We will discuss this important approach later in its own chapter.

At first glance, it seems that the sequence spaces are subspaces of the phylogenetic space, but this is not true: Consider the two words $v = (ab)^d$ and $w = (ba)^d$; then $\rho_L(v,w) = 2$ but $\rho_H(v,w) = 2d$.

To extend the Hamming distance to a metric for all words we may proceed in the following way: Let $A$ be a set of letters. Add a "dummy" letter "-" to $A$. We define a map

$$cl : (A \cup \{-\})^\star \to A^\star \tag{5.93}$$

deleting all dummies in a word from $(A \cup \{-\})^\star$. Then for two words $w$ and $w'$ in $A^\star$ we define the extended Hamming-distance as

$$\begin{aligned}\rho(w,w') \quad &= \min\{\rho_H(\underline{w}, \underline{w'}) : \quad \underline{w}, \underline{w'} \in (A \cup \{-\})^\star, |\underline{w}| = |\underline{w'}|, \\ &\qquad\qquad\qquad\qquad cl(\underline{w}) = w, cl(\underline{w'}) = w'\}.\end{aligned} \tag{5.94}$$

**Observation 5.10.5** *The extended Hamming-distance coincides with the Levenshtein metric.*

As exercise determine or estimate the following quantities:

(a) The number of words in a bounded set in $(A^\star, \rho_L)$.

(b) The diameter of $(A^{\leq n}, \rho_L)$.

# 6

# MARKOV PROCESSES

A Markov chain describes a chance process in which the future state can be predicted from its present state as accurately as if its entire earlier history was known.

## 6.1 TRANSITIONS

Let $\mathcal{S}$ be a finite set of states. Without loss of generality we assume

$$\mathcal{S} = \{1, 2, \ldots, n\}. \tag{6.1}$$

We consider diagrams between states, where the transition from state $i$ to state $j$ occurs with given probability $\alpha_{ij}$, altogether written in a transition matrix

$$A = (\alpha_{ij})_{i,j=1,\ldots,n}. \tag{6.2}$$

Of course, a transition matrix has the following properties:

(i)

$$\alpha_{ij} \geq 0, \tag{6.3}$$

for any $i, j = 1, \ldots, n$.

(ii)

$$\sum_{j=1}^{n} \alpha_{ij} = 1 \tag{6.4}$$

for any $i = 1, \ldots, n$.

Under these conditions, such a matrix is sometimes called stochastic.
The pair $(\mathcal{S}, A)$ is called a Markov process.

Consider

$$A^2 = (\alpha_{ij}^{(2)})_{i,j=1,\ldots,n}. \tag{6.5}$$

Then

$$\begin{aligned}\alpha_{ij}^{(2)} &= \sum_{k=1}^{n} \alpha_{ik}\alpha_{kj} \\ &= \sum_{k=1}^{n} \text{probability for a transition from state } i \text{ to state } k \\ &\quad \star \text{ probability for a transition from state } k \text{ to state } j.\end{aligned}$$

Hence $\alpha_{ij}^{(2)}$ is the probability for a transition from state $i$ to state $j$ in two steps. By induction we find that the probability of the $t$th-step transition of the Markov process is defined as the conditional probability, given the chain is currently in state $i$, that it will be in state $j$ after $t$ additional transitions.

**Theorem 6.1.1** *(Chapman, Kolmogorov)*
*Let $(\mathcal{S}, A)$ be a Markov process. Let*

$$A^t = (\alpha_{ij}^{(t)})_{i,j=1,\ldots,n} \tag{6.6}$$

*be the $t$th power of $A$. Then $\alpha_{ij}^{(t)}$ is the probability for a transition from state $i$ to state $j$ in $t$ steps.*

This means, the $t$th step transition probability matrix may be obtained by multiplying the matrix $A$ by itself $t$ times.
In view of these facts we are interested in $\lim_{t\to\infty} A^t$.

For more information about computational aspects of Markov processes compare [69].

## 6.2 TWO-STATES PROCESSES

As specific case consider $(\{1, 2\}, A)$ with

$$A = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \tag{6.7}$$

where $0 \leq p, q \leq 1$.
For $p = q = 0$ we have $\lim_{t\to\infty} A^t = E$; for $p = q = 1$ the quantity $\lim_{t\to\infty} A^t$ does not exist.
Now we assume $0 < p, q < 1$. We have

$$A = E + \begin{pmatrix} -p & p \\ q & -q \end{pmatrix} = E + B. \tag{6.8}$$

It is easy to see that

$$B^2 = -(p+q) \cdot B,$$

such that

$$B^i = (-1)^{i-1}(p+q)^{i-1} \cdot B, \tag{6.9}$$

for all $i \geq 2$. Then

$$\begin{aligned} A^t &= (E+B)^t \\ &= \sum_{i=0}^{t} \binom{t}{i} B^i \quad \text{by 2.2.4} \\ &= E + \sum_{i=1}^{t} \binom{t}{i} B^i \\ &= E + \sum_{i=1}^{t} \binom{t}{i} (-1)^{i-1}(p+q)^{i-1} B \quad \text{by (6.9)} \\ &= E - \frac{1}{p+q} \sum_{i=1}^{t} \binom{t}{i} (-1)^i (p+q)^i B \\ &= E + \frac{1}{p+q} B - \frac{1}{p+q} \sum_{i=0}^{t} \binom{t}{i} (-1)^i (p+q)^i B \\ &= E + \frac{1}{p+q} B - \frac{(1-p-q)^t}{p+q} B \quad \text{by 2.2.4.} \end{aligned}$$

Since $\lim_{t\to\infty}(1-p-q)^t = 0$ we get

$$\lim_{t\to\infty} A^t = E + \frac{1}{p+q} B. \tag{6.10}$$

**Theorem 6.2.1** *Consider a two-state Markov process. Then*

$$\lim_{t\to\infty} \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}^t = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}. \tag{6.11}$$

## 6.3 EVOLUTIONARY MODELS

Evolutionary models describe the substitution processes in DNA, RNA and amino acid sequences through the time. For simplicity, we will concentrate on DNA sequences, that means the corresponding matrices of the transition probabilities is given by

$$P(t) = \begin{pmatrix} p_{aa}(t) & p_{ac}(t) & p_{ag}(t) & p_{at}(t) \\ p_{ca}(t) & p_{cc}(t) & p_{cg}(t) & p_{ct}(t) \\ p_{ga}(t) & p_{gc}(t) & p_{gg}(t) & p_{gt}(t) \\ p_{ta}(t) & p_{tc}(t) & p_{tg}(t) & p_{tt}(t) \end{pmatrix}, \tag{6.12}$$

for each time parameter $t \geq 0$ the matrix $P(t)$ is a stochastic ones.
Modelling we assume that $P(t)$ gives the probability of all possible states changing in time $t$. We get an continuous-time Markov process. Then we have 6.1.1:

**Theorem 6.3.1** *Let $P(t)$ be the matrix for the transition probabilities.*

$$P(t + t') = P(t) \cdot P(t'). \tag{6.13}$$

Now we assume that such continuous-time Markov processes which are differentiable at every $t \geq 0$. Then it follows for $h > 0$:

$$\begin{aligned} \frac{P(t+h) - P(t)}{h} &= \frac{P(t)P(h) - P(t)}{h} \quad \text{in view of 6.3.1} \\ &= \frac{P(t)(P(h) - E)}{h} \\ &= P(t) \cdot \frac{P(h) - P(0)}{h}. \end{aligned}$$

When $h \to 0$ this identity implies

$$P'(t) = P(t) \cdot P'(0). \tag{6.14}$$

This differential equation gives the following solution.

**Theorem 6.3.2** *Under the assumptions given above the matrix $P(t)$ has the form*

$$P(t) = e^{tQ}, \tag{6.15}$$

*where $Q$ is some (fixed) matrix.*

Recall, that for a square matrix $A$ we define the exponential matrix $e^A$ by the sum of the following series:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

The matrix $Q$ is called the matrix of instantaneous change or the rate matrix. It has the following important properties:

(a) It holds the "inverse" identity

$$Q = P'(0). \tag{6.16}$$

(b) The elements in each row of $Q$ sum up to 0. In particular,

$$\det Q = 0. \tag{6.17}$$

By varying the matrix $Q$ one obtains several models:

The Jukes-Cantor model is the oldest model and assumes that the probabilities to find a nucleotide site are equal in any of the four possible states and for all time $t$.[1] The matrix of instantaneous change is given by setting

$$Q = \frac{1}{4} \begin{pmatrix} -3\alpha & \alpha & \alpha & \alpha \\ \alpha & -3\alpha & \alpha & \alpha \\ \alpha & \alpha & -3\alpha & \alpha \\ \alpha & \alpha & \alpha & -3\alpha \end{pmatrix}, \tag{6.18}$$

where $\alpha$ is a positive real number, called the evolutionary rate.

We will calculate the corresponding matrix $P(t)$.
First, by induction, it is easy to see that

$$Q^n = (-\alpha)^{n-1} Q \tag{6.19}$$

is true for all integers $n \geq 1$. Now we find $P(t) = \exp(tQ)$ by the following calculations.

$$\begin{aligned} P(t) &= \sum_{n=0}^{\infty} \frac{t^n Q^n}{n!} \\ &= E + \sum_{n=1}^{\infty} \frac{t^n Q^n}{n!} \end{aligned}$$

[1] This assumption is not very realistic.

$$\begin{aligned} &= E + \left( \sum_{n=1}^{\infty} \frac{t^n(-\alpha)^{n-1}}{n!} \right) Q \quad \text{in view of (6.19)} \\ &= E - \frac{1}{\alpha} \left( \sum_{n=1}^{\infty} \frac{(-t\alpha)^n}{n!} \right) Q \\ &= E - \frac{1}{\alpha} \left( e^{-t\alpha} - 1 \right) Q. \end{aligned}$$

This implies

**Theorem 6.3.3** *The transition matrix in the Jukes-Cantor model equals*

$$p_{ij}(t) = \begin{cases} \frac{1}{4} + \frac{3}{4} e^{-t\alpha} & : \ i = j \\ \frac{1}{4} - \frac{1}{4} e^{-t\alpha} & : \ i \neq j \end{cases}$$

*($i, j \in \{a, c, g, t\}$)*

The Kimura model models a certain difference between two types of nucleotide substitutions: Purines into pyrimidines or vice versa; and inside purines or pyrimidines. It is given by setting

$$Q = \frac{1}{4} \begin{pmatrix} -(2\beta+1)\alpha & \beta\alpha & \alpha & \beta\alpha \\ \beta\alpha & -(2\beta+1)\alpha & \beta\alpha & \alpha \\ \alpha & \beta\alpha & -(2\beta+1)\alpha & \beta\alpha \\ \beta\alpha & \alpha & \beta\alpha & -(2\beta+1)\alpha \end{pmatrix}, \qquad (6.20)$$

with two parameters $\alpha, \beta > 0$.

As before, we find the following result.

**Theorem 6.3.4** *The transition matrix in the Kimura model equals*

$$p_{ij}(t) = \begin{cases} \frac{1}{4} + \frac{1}{4} e^{-t\alpha\beta} - \frac{1}{2} e^{-t\alpha(\beta+1)/2} & : \ (i,j) = (a,g), (g,a), (c,t) \ \textit{or} \ (t,c) \\ \frac{1}{4} - \frac{1}{4} e^{-t\alpha\beta} & : \ (i,j) = (a,c), (c,a), (a,t), (t,a), \\ & : \ (c,g), (g,c), (g,t) \ \textit{or} \ (t,g) \\ \frac{1}{4} + \frac{1}{4} e^{-t\alpha\beta} + \frac{1}{2} e^{-t\alpha(\beta+1)/2} & : \ \textit{otherwise} \end{cases}$$

For more information, and other models, compare [46].

# 7

# SIMILARITY OF WORDS

> Einstein said: "God does not play dice." He was right. God plays scrabble.
>
> Philip Gold

In the biological context the equality of words makes no sense, since mutations do not allow identical sequences in reality. On the other hand, in biomolecular sequences, high sequence similarity usually implies significant functional and structural similarity.[1]

Let $A$ be an alphabet. We consider the set $A^\star$ of all words over $A$. Our interest is to define measures on $A^\star$ which reflect the "proximity" of two words. Here, two different approaches are to be distinguished: distance and similarity. We will introduce both measures in the greatest possible generality. This is necessary, since evolution, as reflected at the molecular level, proceeds by a series of insertions, deletions and substitutions of letters, as well as other far rarer mechanisms which we are ignore here, since we observe not complete genomes, only genes or other "smaller" words.

[1] But note that the converse is, in general, not true. And in reality, for applications in biology it is sometimes necessary to take into account several other properties of the macro-molecules to measure their similarity, for instance structure, expression and pathway similarity, compare [49].

## 7.1 DISTANCES BETWEEN WORDS

To find a metric for words over $A$ we consider a cost measure $(c, h)$ for the letters given by

- A function $c : A \times A \to \mathbb{R}_{\geq 0}$, which satisfies the following conditions:
  - (i) $c$ is non-negative: $c(a, b) \geq 0$;
  - (ii) $c(a, a) = 0$; and
  - (iii) $c$ is symmetric: $c(a, b) = c(b, a)$ for any $a, b \in A$.
- A positive real number $h$.

The substitution of a letter $b$ for a letter $a$ costs $c(b, a) = c(a, b)$. The insertion or deletion of a letter effectively transforms a non-gap letter in one word to a gap in the other. Since we do not know the direction of the change through time, it is useful to group both operations under the term indel. Each indel costs $h$.
The distance $\rho(w, w')$, between two sequences $w, w' \in A^{\star}$ according to a cost measure is the minimum of the costs running over all series of operations transforming $w$ into $w'$.

**Observation 7.1.1** *The function $\rho$ is a pseudo-metric. If, moreover, the function $c$ satisfies the non-degeneracy property, i.e. that $c(a, b) = 0$ holds if and only if $a = b$, then $\rho$ is a metric.*

Note, that we do not assume that $c$ satiesfies the triangle inequality, but we can assume this. The reason for this assumption is that even if we start with a cost measure $(c, h)$ that does not satisfy it, we can always define a new pair $(c', h)$ that does satisfy it and produces the same metric. Namely, if three letters $a_1, a_2$ and $a_3$ are such that $c(a_1, a_2) > c(a_1, a_3) + c(a_3, a_2)$, then every time we need to replace $a_1$ by $a_2$ we will not do it directly but rather replace $a_1$ by $a_3$ and later $a_3$ by $a_2$, producing the same effect at a lower cost. Moreover, using the the same reasoning, the restriction of the metric $\rho$ to the alphabet itself need not be $c$. This is only true if the function $c$ satisfies the triangle inequality.

An example: For the cost measure $(c, h)$ defined by

| | a | c | g | t |
|---|---|---|---|---|
| a | 0 | 2 | 1 | 2 |
| c | | 0 | 2 | 1 |
| g | | | 0 | 2 |
| t | | | | 0 |

and $h = 4$, we find $\rho(agc, a^3c) = 5$, $\rho(acg, a^3c) = 7$ and

$$\rho((ag)^d, (ga)^d) = \begin{cases} 2d & \text{if} \quad d = 1, 2, 3 \\ 8 & \text{if} \quad d \geq 4 \end{cases}$$

## 7.2 SIMILARITY OF WORDS

Another approach uses similarity. The procedure used to find such quantity is called sequence alignment and depends on a scoring system. The elongated sequences in an alignment should be as similar as possible according to some predefined scoring system.

### 7.2.1 Alignments

Sequence alignment is the identification of residue-residue correspondences. It is **the** basic tool of bioinformatics.

Any assignment of correspondences that preserves the order of the residues within sequences is an alignment; gaps may be introduced: Given two sequences $w$ and $w'$ over the same alphabet, an alignment of $w$ and $w'$ is a matrix with the following properties:

(i) There are two rows, each row for the elongated sequence for $w$ and $w'$, which preserves the left-to-right ordering of the letters, but uses dummy symbols;

(ii) The elongated sequences are of the same length;

(iii) There is no column for which the elongated sequences both have a dummy.

For instance consider the two words $w = gt^2a^2c^2$ and $w' = gatc$. The following arrays are all alignments for $w$ and $w'$:

```
g t t a a c c
g a t c - - -
```

```
g t t a a c c - - - -
- - - - - - - g a t c
```

and

```
g t t a a c c
g - - a t c -
```

where ”-” denotes a ”dummy” symbol.

The conditions for a pairwise alignment implies

$$\max\{|w|, |w'|\} \leq l \leq |w| + |w'|, \tag{7.1}$$

where $l$ denotes the common length of the elongated sequences. Consequently, there are a finite number of alignments for a given pair of sequences. More exactly, there are

$$\binom{n+m}{n} = \binom{n+m}{m} \tag{7.2}$$

alignments of two sequences with $n$ and $m$ letters, respectively. For a proof see [88].

### 7.2.2 Multiple alignments

In the context of molecular biology, multiple sequence comparison is the most critical cutting-edge tool for extracting and showing biologically important factors that a set of sequences has in common. It plays an essential role in two related areas:

- Finding highly conserved subregions among a collection of sequences; and
- Inferring the evolutionary history of some species from their associated sequences.

One central technique for multiple sequence comparison involves multiple alignment. Here, a (global) multiple alignment of $n > 2$ sequences $w_1, \ldots, w_n$ is a natural generalization of the alignment of two sequences. That means that we insert dummies into, or at either end of, each of the sequences to produce a new collection of elongated sequences that obeys these rules:

(i) All elongated sequences have the same length, $l$;

(ii) There is no position at which all the elongated sequences have a dummy.

Then the sequences are arrayed in a matrix of $n$ rows and $l$ columns, where

$$\max_{i=1,\ldots,n} |w_i| \leq l \leq \sum_{i=1}^{n} |w_i|. \tag{7.3}$$

Consequently, there are a finite number of multiple alignments for a collection of sequences.

In any case, the alignment array can be summarized in a single sequence called a consensus sequence, which is frequently added at the end of the alignment. It is common in computational molecular biology to compute a multiple alignment for a set of sequences, and then represent those sequences by the consensus sequence derived from the alignment.
The consensus sequence consists of letters that summarizes the letters of the alignment in each column. A simple way to calculate a consensus sequence is to use the so-called majority rule (MR), which chooses the most frequently occuring letter in each column. We distinguish between two rules:

- The normal rule uses the alphabet $A \cup \{-\}$.
- The restricted rule uses only the alphabet $A$.

An example compares the word for SCHOOL in different languages:

| Language | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| German | - | S | C | H | U | - | L | E |
| English | - | S | C | H | O | O | L | - |
| French | E | - | C | - | O | - | L | E |
| Italian | - | S | C | - | U | O | L | A |
| Consensus, MR | - | S | C | H or - | O or U | O or - | L | E |
| Consensus, restricted MR | E | S | C | H | O or U | O | L | E |

### 7.2.3 Scoring systems

Given an alignment between two sequences, we assign a score to it as follows: Each column of the alignment will receive a certain value depending on its contents and the total score for the alignment will be the sum of the values assigned to its columns.
Let an alignment between two sequences be given. We use the following abbreviations:

(a) If a column has two identical symbols we will call it a match;

(b) A column with two different symbols is called a mismatch; and

(c) A column with a dummy in one row is called a gap.

We use a scoring system $(p, g)$, which is given by

- A symmetric function $p : A \times A \to \mathbb{R}$, and
- A non-positive real number $g$.

The array of $p$ is called the (substitution) score matrix. The value $p(a, b)$ scores pairs of aligned letters $a$ and $b$. The penalty $g$ is used to penalize gaps. In general, we assume that $p(a, a) > 0$, for $a \in A$, and $g < 0$.[2] Clearly, the selection of an appropriate score matrix is crucial for achieving "good" alignments.
A scoring system assigns a value, called the score, to each possible alignment by adding the values for each column.
The similarity $\mathrm{sim}(w, w')$, between two sequences $w, w' \in A^\star$ according to a scoring system is the maximum of the scores running over all alignments of $w$ and $w'$.

As an example use the following score matrix:

| | a | c | g | t |
|---|---|---|---|---|
| a | 4 | 1 | 2 | 1 |
| c | | 4 | 1 | 2 |
| g | | | 4 | 1 |
| t | | | | 4 |

[2] And unlikely substitutions are penalized with a negative score.

and gap-penalty 0 for our alignments above:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| g | t | t | a | a | c | c | |
| g | a | t | c | - | - | - | |
| 4 | +1 | +4 | +1 | +0 | +0 | +0 | =10 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| g | t | t | a | a | c | c | - | - | - | - | |
| - | - | - | - | - | - | - | g | a | t | c | |
| 0 | +0 | +0 | +0 | +0 | +0 | +0 | +0 | +0 | +0 | +0 | = 0 |

and

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| g | t | t | a | a | c | c | |
| g | - | - | a | t | c | - | |
| 4 | +0 | +0 | +4 | +1 | +4 | +0 | =13 |

There are different manners in which a (substitution) score matrix can be derived. In general, in a biological context a scoring matrix $p$ is a table of values that describe the probability of a residue (amino acid or base) pair occuring in an alignment. The approach is good, if the score matrix produces good alignments.

Substitution matrices for amino acids are complicated because they reflect the chemical nature and the frequency of occurrence of the amino acids, see [4]. The PAM (Point Accepted Mutation) series of score matrices are frequently used for protein alignments [2] and [27]. Each entry in a PAM matrix gives the logarithm of the ratio of the frequency at which a pair of residues is observed in pairwise comparisons of homologous proteins to the frequency expected due to chance alone. Amino acids that regularly replace each other have a positive score, while amino acids that rarely replace each other have a negative score.

Substitution matrices for bases in DNA or RNA sequences are simpler: in most cases, it is reasonable to assume that a:t and g:c occur in roughly equal proportions.

## 7.3 THE INTERRELATION BETWEEN DISTANCE AND SIMILARITY

The concepts of distance and of similarity are essentially dual. We will describe what this mean.

**Algorithm 7.3.1** *Given a cost measure $(c, h)$ and a constant $K$, we can define a scoring system $(p, g)$ as follows:*

$$\begin{aligned} p(a,b) &= K - c(a,b), \\ g &= -h + \frac{K}{2}, \end{aligned}$$

*under the constraint*

$$K \leq 2h. \tag{7.4}$$

*And conversely, given a scoring system $(p, g)$ with the property that $p(a,a) = K$ for all $a \in A$, we can define a cost measure $(c, h)$ as follows:*

$$\begin{aligned} c(a,b) &= K - p(a,b), \\ h &= -g + \frac{K}{2}, \end{aligned}$$

*under the constraints*

$$\begin{aligned} K &\geq \max\{p(a,b) : a, b \in A\}, \ \textit{and} \\ K &> 2g. \end{aligned}$$

In other words, we have the following interrelation between a cost measure $(c, h)$ and a scoring system $(p, g)$:

$$p(a,b) - 2g = 2h - c(a,b) \tag{7.5}$$

for all $a, b \in A$, which obviously reflects the duality. Roughly speaking, "large distance" is "small similarity" and vice versa. Moreover, distance computation can be reduced to similarity computation:

**Theorem 7.3.2** *(Smith, Waterman, Fitch [77], Setubal, Meidanis [76], Waterman [87]) A cost measure and the corresponding scoring system as in 7.3.1 are given for a certain value $K$. Let $w$ and $w'$ be sequences (words) over $A$. Then*

$$\rho(w, w') + sim(w, w') = \frac{K}{2} \cdot (|w| + |w'|). \tag{7.6}$$

*Both the cost measure and the corresponding scoring system yield the same optimal alignments.*[3]

*Sketch of the proof.* Let $w$ and $w'$ be sequences of length $m$ and $n$ respectively, and let $\alpha$ be an alignment between $w$ and $w'$. We define a series $\sigma$ of operations transforming $w$ into $w'$ by dividing $\alpha$ into columns corresponding to the operations in a natural way:

- matches and mismatches in the alignment correspond to substitutions in the transformation;
- gaps in the alignment corresponds to indels in the transformation.

We shall now compute the score of $\alpha$ and the cost of $\sigma$. Suppose there are exactly $l$ letters which are matched or mismatched in $\alpha$, occupying positions $w_i$ in $w$ and $w'_i$ in $w'$, $1 \leq i \leq l$. Suppose further that there are exactly $r$ gaps in $\alpha$. Then

$$\text{score}(\alpha) = \sum_{i=1}^{l} p(w_i, w'_i) + rg. \tag{7.7}$$

On the other hand, the cost of $\sigma$ is

$$\text{cost}(\sigma) = \sum_{i=1}^{l} c(w_i, w'_i) + rh. \tag{7.8}$$

Memberwise addition of (7.7) and (7.8) in conjunction with 7.3.1 give

$$\text{score}(\alpha) + \text{cost}(\sigma) = lK + r\frac{K}{2}. \tag{7.9}$$

Moreover the values of $l$ and $r$ are not independent: each match and mismatch uses two letters and each gap uses one. Therefore, the total number of letters must be

$$m + n = 2l + r. \tag{7.10}$$

Then (7.9) can be written as

$$\text{score}(\alpha) + \text{cost}(\sigma) = \frac{K}{2} \cdot (m + n). \tag{7.11}$$

Since this is true for any alignment, we have one half of the assertion.
The other half follows similarly.

[3] Although with different scores. But using the formula given in 7.3.1 the distance is the same.

□

All these considerations imply that, from the mathematical standpoint, an alignment and an edit transformation are equivalent ways to describe a relationship between two sequences. But we should note what Gusfield [41] wrote:

> Although an alignment and an edit transcript are mathematically equivalent, from a modeling standpoint, an edit transcript is quite different from an alignment. An edit transcript emphasizes the putative *mutational events* (point mutations in the model so far) that transform one string to another, whereas an alignment only displays a relationship between two strings. The distinction is one of *process* versus *product*. Different evolutionary models are formalized via different permitted string operations, and yet these can result in the same alignment. So an alignment alone blurs the mathematical model. This is often a pedantic point but proves helpful in some discussions of evolutionary modeling.

The similarity-based approach is more general than that of distance, since

- The distance-based approach is restricted to global comparisons only, it is not suitable for local ones. For an algorithm and derivations of our basic technique compare [76].
- With similarities we can penalize gaps depending on their lengths. This cannot be done with metrics. This is an important observation, since if two aligned sequences are for functional protein coding genes, then any gaps would be expected to have lengths that were multiples of three, to preserve the reading frame of the gene; and for ribosomal genes there may be aspects of the secondary structure that can be used to evaluate the plausibility of the various gaps introduced in an alignment.
- In any case we assume that for a cost measure $(c, h)$ the equality $c(a, a) = 0$ holds for all letters $a$. On the other hand, there are scoring systems $(p, g)$ conceivable in which for different letters $a$ and $b$ we have $p(a, a) \neq p(b, b)$.[4]
- For a generalized scoring system, derived dissimilarity need not satisfy the triangle inequality.

[4] Particularly, the PAM matrices.

## 7.4 SPECIFIC CASES

A simplified scoring system, called a match-mismatch-gap system, is given if all matches have the same value $M = p(a,a)$ and likewise all mismatches have the same value $m = p(a,b)$, $a \neq b$. Of course, we assume that $M \geq 0$ and $g \leq 0$. Additionally, a substitution $(a,b)$ must be "cheaper" than two indels $(a-, -b)$. Hence, we have

**Corollary 7.4.1** *Let $(M,m,g)$ be a scoring system with only values for matches, mismatches and gaps. Then a cost measure $(c,h)$ having $c(a,a) = 0$ and $c(a,b) = c > 0$ is given by*

$$\begin{aligned} c &= M - m, \\ h &= \frac{M}{2} - g, \end{aligned}$$

*provided that*

$$M \geq m \geq 2g, \tag{7.12}$$

*in which at least one inequality is strict, $M \geq 0$, and $g \leq 0$.*

As examples we consider several standard systems:

(i) The Levenshtein distance, that is $c = 1$ and $h = 1$. We may choose match score $M = 2$, mismatch score $m = 1$ and gap score $g = 0$.

(ii) If we wish to measure the distance by

$$\rho(w,w') = \#\text{ substitutions } + h \cdot \#\text{ indels}, \tag{7.13}$$

for $h \geq 1$ (i.e. that gaps are $h$ times as costly as substitutions), we may choose $M = 2$, $m = 1$ and $g = 1 - h$.

(iii) A "normed" match-mismatch-gap system with one free parameter is given by $(1, m, 0)$ where $1 \geq m \geq 0$. Equivalently, we have a cost measure with $c = 1 - m$ and $h = 1/2$. In particular, we consider

**The problem of longest common subsequence**
**Given:** A set of sequences over the same alphabet.
**Find:** A longest sequence contained in each of the given sequences.

The search for a pair of words uses the match-mismatch-gap system (1,0,0) which implies $c = 1$ and $h = 1/2$.

## 7.5 THE ALGORITHM

How can we find the similarity of or the distance between two words? Clearly, the consideration of all possible alignments does not make sense, since there are too many; see (7.2). Observe that we cannot change the order of the letters in the words. This fact suggests that a dynamic programming approach will be useful, which finds the solution by first breaking the original problem into smaller subproblems and then solving all these subproblems, storing each intermediate solution in a table along with a score, and finally choosing the sequence of solutions that yields the highest score.
Let $w$ and $w'$ be two sequences over $A$ with length $m$ and $n$, respectively. The algorithms use a $(m+1) \times (n+1)$ matrix, and determine the values of this matrix in the following way:

**Algorithm 7.5.1** *Let $w = a[1]a[2]\dots a[m]$ and $w' = b[1]b[2]\dots b[n]$ be two sequences in $A^\star$, equipped with a scoring system $(p, q)$. Then, we find the similarity $sim(w, w')$ =$sim[m, n]$ by the following procedure.*

```
1. for $i := 0$ to $m$ do
     $sim[i, 0] := i \cdot g$;
2. for $j := 0$ to $n$ do
     $sim[0, j] := j \cdot g$;
3. for $i := 1$ to $m$ do
     for $j := 1$ to $n$ do
     $sim[i, j] := \max\{sim[i-1, j] + g, sim[i-1, j-1] + p[i, j], sim[i, j-1] + g\}$
```

As example we consider the similarity between NAME and MEAN under the match-mismatch-gap system $(4, 1, 0)$.

| | | N | A | M | E |
|---|---|---|---|---|---|
| | 0 | 0 | 0 | 0 | 0 |
| M | 0 | 1 | 1 | 4 | 4 |
| E | 0 | 1 | 2 | 4 | 8 |
| A | 0 | 1 | 5 | 5 | 8 |
| N | 0 | 4 | 5 | 6 | 8 |

Hence $\text{sim}(\text{NAME}, \text{MEAN}) = 8$. By the equivalent cost measure $c = 3$ and $h = 2$ there is $\rho(\text{NAME}, \text{MEAN}) = 8$.

An alignment of two words $w$ and $w'$ is called an optimal alignment if its score equals $\text{sim}(w, w')$. The algorithm, as stated above, only computes the similarity of the words. For the explicit construction of an optimal alignment, the algorithm has to be supplemented by a backtracking procedure. This alignment corresponding to the similarity may well not be unique; but all such alignments can be found "backtracking" from the cell $\text{sim}[m, n]$ to the cell $\text{sim}[0, 0]$ in all possible ways.

In our example above we have

$$\begin{array}{cccccc} \text{N} & \text{A} & \text{M} & \text{E} & \text{-} & \text{-} \\ \text{-} & \text{-} & \text{M} & \text{E} & \text{A} & \text{N} \end{array}$$

Note that this method to determine the similarity of two sequences is relatively fast but still too slow for most practical work, where the length of the sequences and the number of sequences to be compared are very large. This comes from the following often used question: You already have a particular protein or nucleic acid sequence that you are interested in and you need to find other sequences that are related to it.[5]

There are heuristic methods which are more efficiently for "similarity-searching" an entry in a collection of sequences, namely the well-known BLAST method, compare [76].

## 7.6 OPTIMAL MULTIPLE ALIGNMENTS

Although the notation of a multiple alignment is easily extended from two to many sequences, the score or the cost of a multiple alignment is not easily generalized. There is no function that has been universally accepted for multiple alignment as distance or similarity has been for pairwise alignment.

Recall that a cost measure $(c, h)$ for an alphabet $A$ to compare two sequences can be also written as a function $f : (A \cup \{-\})^2 \to \mathbb{R}$, where $-$ is the "dummy"

[5] By "related" we mean that another sequence is sufficiently similar to the sequence of interest that we belive the two sequences share a common ancestor.

symbol, $- \notin A$, and

$$\begin{aligned} f(a,b) &= c(a,b), && (7.14)\\ f(a,-) = f(-,b) &= h. && (7.15)\end{aligned}$$

($f(-,-)$ is not defined.) $A \cup \{-\}$ is called the extended alphabet, and such a function $f$, extended to $n \geq 2$ values, is called a generalized cost measure. More precisely: A generalized cost measure is a function $f : (A \cup \{-\})^n \to I\!R_{\geq 0}$, which satisfies the following conditions:

(i) $f$ is non-negative: $f(a_1, \ldots, a_n) \geq 0$;

(ii) $f(a, \ldots, a) = 0$, for each $a \in A$;
$f(-, \ldots, -)$ is not defined;

(iii) $f(a_1, \ldots, a_n) > 0$ if $a_i = -$ holds for at least one index $i$;

(iv) $f$ is symmetric:

$$f(a_{\pi(1)}, \ldots, a_{\pi(n)}) = f(a_1, \ldots, a_n) \qquad (7.16)$$

holds true for any permutation $\pi$.

With this in mind, we have several methods to find the generalized similarity, see [22], [88] and [85].

# 8

# GRAPHS AND TREES

We have to introduce several knowledge of graphs and networks. Graphs are among the most basic of all mathematical structures. Correspondingly, they have many different versions, representations and incarnations.

## 8.1 GRAPHS

A graph $G$ is defined to be a pair $(V, E)$ where

- $V$ is any finite set of elements, called vertices, and
- $E$ is a finite family of elements which are unordered pairs of vertices, called edges.

The notation $e = \underline{uv}$ means that the edge $e$ joins the vertices $u$ and $v$. In this case, we say that the vertices $u$ and $v$ are incident to this edge and that $u$ and $v$ are the endvertices of $e$. Two vertices $u$ and $v$ are called adjacent in the graph $G$ if $\underline{uv}$ is an edge of $G$.[1]
$N(v) = N_G(v)$ denotes the set of all vertices adjacent to the vertex $v$ and is called the set of all neighbors of $v$ in $G$.
For a vertex $v$ of a graph $G$ the degree $g_G(v)$ is defined as the number of edges which are incident to $v$. If $G$ has no parallel edges then the cardinality of $N(v) = N_G(v)$ is the degree of the vertex:

$$g(v) = g_G(v) = |N_G(v)|. \tag{8.1}$$

[1] In any case, we assume that $u \neq v$, that means we do not admit loops.

If we sum up all the vertex degrees in a graph, we count each edge exactly twice, once from each of its endvertices. Thus,

**Observation 8.1.1** *In any graph $G = (V, E)$ the equality*

$$\sum_{v \in V} g_G(v) = 2 \cdot |E| \tag{8.2}$$

*holds. Particularly, in every graph the number of vertices with odd degree is even.*

A graph $G$ is said to be a complete graph if any two vertices are adjacent. A complete graph with $n$ vertices has exactly

$$\binom{n}{2} = \frac{n(n-1)}{2} \tag{8.3}$$

edges.

Let $G = (V, E)$ be a graph. Then $G' = (V', E')$ is called a subgraph of $G$ if $V'$ is a subset of $V$ and $E'$ is a subset of $E$ such that any edge in $E'$ joins vertices from $V'$. In other terms,

$$V' \subseteq V \tag{8.4}$$

and

$$E' \subseteq E \cap \binom{V'}{2}. \tag{8.5}$$

A chain is a sequence $v_1, e_1, v_2, e_2, v_3, ..., v_m, e_m, v_{m+1}$ of edges and vertices of $G$ such that the edge $e_i$ is incident to the vertices $v_i$ and $v_{i+1}$ for any index $i = 1, ..., m$. A chain in which each vertex appears at most once is called a path; more exactly, the path interconnecting the vertices $v_1$ and $v_{m+1}$. Then the number $m$ denotes the length of the path. A single vertex is a path of length 0.

A cycle is a chain with at least one edge and with the following properties: No edge appears twice in the sequence and the two endvertices of the chain are the same. A graph which does not contain a cycle is called acyclic.

A key notion in graph theory is that of a connected graph.[2] It is intuitively clear what this should mean, and it is also easy to formulate this property: A

[2] In a natural sense, graph theory is the study of connectivity.

graph $G = (V, E)$ is called a connected graph if for any two vertices there is a path (or, equivalently, a chain) interconnecting them. Let $G = (V, E)$ be a graph and let $v$ and $v'$ be two vertices of $G$. Clearly,

**Observation 8.1.2** *The relation "There is a path in $G$ connecting $v$ and $v'$" is an equivalence relation on $V \times V$.*

The equivalence classes of this relation divide $V$ into subsets, which create connected subgraphs of $G$. These classes are called the connected components, or briefly the components of the graph $G$. A component is a maximal subgraph that is connected. A connected graph has exactly one component.

## 8.2 THE METRIC CLOSURE OF A NETWORK

We consider networks. These are (connected) graphs $G = (V, E)$ equipped with a length function $f : E \to I\!R$. This function on the edges of $G$ is constrained to take only strictly positive values.[3]
The simplest question, which will be of great importance in further considerations, is to look for the interconnecting chains of shortest length between vertices in the network:

**The Shortest Path Problem**
**Given:** A network $G = (V, E, f)$ and two vertices $v$ and $v'$ of $G$.
**Find:** A path connecting $v$ and $v'$ with minimal length.

A solution is called a shortest path (between the vertices $v$ and $v'$ in $G$).

**Observation 8.2.1** *Let $G = (V, E, f)$ be a network. Define the function $\rho$ on $V \times V$ so that*

$$\rho(v, v') = \textit{the length of a shortest path between the vertices } v \textit{ and } v' \textit{ in } G, \tag{8.6}$$

*for two vertices $v$ and $v'$. Then $(V, \rho)$ is a metric space.*

[3] Nevertheless saying it explicitely, sometimes we will use a length function which has the value 0 for several edges.

The space $(V, \rho)$ is called the metric closure $G^f$ of the network $G = (V, E, f)$. We can also define $G^f$ as the complete graph on $V$ such that the length of an edge $\underline{vv'}$ in $G^f$ is the length of a shortest path between $v$ and $v'$ in $G$. Note that $G$ is a subgraph of $G^f$, but the restriction of $\rho$ on $G$ must not be $f$.

The problem of finding shortest paths in a network is easy to solve by the dynamic programming technique. More precisely, we use the following observation, called Bellman's principle of optimality:

**Observation 8.2.2** *(Bellman [8]) Let $G = (V, E, f)$ be a network, and let $v$ and $v'$ be two vertices of $G$. If $e = \underline{wv'}$ is the final edge of some shortest path $v, \ldots, w, v'$ from $v$ to $v'$, then $v, \ldots, w$ (that is the path without the edge $e$) is a shortest path from $v$ to $w$.*

Roughly speaking: An optimal strategy contains only optimal substrategies. The observation gives immediately

**Algorithm 8.2.3** *(Dijkstra [28]) Let $G = (V, E, f)$ be a network. A shortest path between the vertices $v$ and $v'$ can be found by the following procedure:*

1. *Start with the vertex $v$;*
   *Label $v$ with 0: $L(v) := 0$; all other vertices are unlabelled;*
2. *Determine $\min\{L(v_1) + f(\underline{v_1 v_2})\}$ where $v_1$ and $v_2$ are adjacent vertices, $v_1$ labelled and $v_2$ not;*
   *Choose $\tilde{v}_1$ and $\tilde{v}_2$ which attain the minimum;*
   *Label $\tilde{v}_2$ by $L(\tilde{v}_2) = L(\tilde{v}_1) + f(\underline{\tilde{v}_1 \tilde{v}_2})$;*
3. *Repeat the second step until $v'$ is labelled.*

*For all labelled vertices $w$ the quantity $L(w)$ is the length of a shortest path connecting $v$ and $w$:*

$$\rho(v, w) = L(w).$$

Now it is easy to construct the metric closure $G^f$: it is sufficient to apply 8.2.3 $|V|$ times.

## 8.3 TREES

A tree is defined to be a connected graph without cycles. A forest is defined as a graph whose connected components are trees. That means a forest is a acyclic graph.

A vertex with degree one is called a leaf. It is easy to see that each tree with more than one vertex has at least two leaves. A vertex in a tree that is not a leaf is called an internal vertex.

**Observation 8.3.1** *Let $G = (V, E)$ be a graph with $n$ vertices, where $n > 1$.*[4] *Then the following properties are pairwise equivalent (and each characterizes a tree):*

1. *$G$ is connected and has no cycles.*
2. *$G$ is connected and contains exactly $n - 1$ edges.*
3. *$G$ has exactly $n - 1$ edges and has no cycles.*
4. *$G$ is maximally acyclic; that means $G$ has no cycles, and if a new edge is added to $G$, exactly one cycle is created.*
5. *$G$ is minimally connected; that means $G$ is connected, and if any edge is removed, the remaining graph is not connected.*
6. *Each pair of vertices of $G$ is connected by exactly one path.*

The proof is intuitively clear.

As a consequence of our considerations, we consider a tree $T = (V, E)$ with $n$ vertices. Let $n_i$ be the number of vertices of degree $i$ and $\Delta = \Delta(T)$ is the maximum degree in the tree $T$. Then, of course,

$$n_1 + n_2 + \ldots + n_\Delta = n. \tag{8.7}$$

In view of 8.1.1 and 8.3.1, we have

$$n_1 + 2 \cdot n_2 + \ldots + \Delta \cdot n_\Delta = 2|E| = 2n - 2. \tag{8.8}$$

When we subtract this equation from twice of (8.7) we get

[4] By definition a graph with one vertex and without edges is also a tree.

**Observation 8.3.2** *It holds*

$$n_1 = 2 + \sum_{i=3}^{\Delta(T)} (i-2) \cdot n_i, \tag{8.9}$$

*for any tree, whereby $n_i$ denotes the number of vertices of degree $i$ and $\Delta(T)$ is the maximum degree in the tree.*

## 8.4 MINIMUM SPANNING TREES

Let $G = (V, E)$ be a graph. A subgraph $G' = (V, E')$ is called a spanning tree of $G$ if $G'$ is a tree. If $G'$ is a spanning tree of $G$, then $G$ itself must be connected. Conversely, if $G = (V, E)$ is a connected graph, then $G$ contains a subgraph $G' = (V, E')$ minimal with respect to the property that $G'$ is connected. The graph $G'$ is a spanning tree of $G$. Hence,

**Observation 8.4.1** *A graph is connected if and only if it contains a spanning tree.*

Additionally, we assume that a function $f : E \to \mathbb{R}$ is given for the edges of the graph $G$; we consider a network $G = (V, E, f)$. Then we define the length of a subgraph $G' = (V, E')$ of the graph $G$ as

$$L(G') := \sum_{e \in E'} f(e). \tag{8.10}$$

**The Minimum spanning tree problem**
**Given:** A network $G = (V, E, f)$.
**Find:** A spanning tree $T = (V, E')$, $E' \subseteq E$, which minimizes the length $L(T)$.

A solution is called a minimum spanning tree for the network $G$.

The most recently discovered of the classical algorithms is that of Kruskal, created:

**Algorithm 8.4.2** *(Kruskal [55]) Given a connected graph $G = (V, E)$ with a length-function $f : E \to I\!R$, a minimum spanning tree $T$ for $G$ can be found by the following procedure:*

1. *Start with the forest $T = (V, \emptyset)$;*
2. *Sequentially choose the shortest edge that does not form a cycle with edges already chosen;*
3. *Stop when all vertices are connected, that is when $|V| - 1$ edges have been chosen.*

## 8.5 LABELLED AND SEMI-LABELLED TREES

We have to distinguish between labelled an unlabelled trees. A tree $T = (V, E)$ with $n$ vertices is called labelled if a bijective mapping from $V$ onto the set $\{1, \ldots, n\}$ of integers is given.[5] On the other hand, in the case of unlabelled trees the word "different" means non-isomorphic, and each set of isomorphic trees is counted as one.

In phylogenetics we search for a tree interconnecting a set $N$ of "living entities" (species, genes, sequences, words - roughly speaking: names). Such a partially labelled tree (semi-labelled tree) is usually called an $N$-tree, which means:

- The tree has exactly $|N|$ leaves, each labelled by a different element of $N$;
- All internal vertices are unlabelled;
- The degree of each internal vertex is at least 3.
- Sometimes we accept an exception, namely that exactly one internal vertex is marked, and is permitted to have degree 2. Then this vertex is called the root of the tree, and such a tree is called a rooted $N$-tree.

[5] Or onto another set of $n$ distinguished names.

### 8.5.1 Counting labelled trees

It is not the purpose of this chapter to provide a complete survey of counting methods for trees. We will focus on the counting of specific classes of trees, which are important in investigations about phylogeny.

We start counting with the number of different labelled trees and we will describe this number in terms of the vertex degrees.
Let $T = (V, E)$ be a tree with $n$ vertices $v_1, ..., v_n$, and let $g_i = g(v_i)$ be the degree of each vertex $v_i$. Then, obviously, each of the numbers $g_i$ is a positive integer, and, in view of 8.1.1 and 8.3.1,

$$\sum_{i=1}^{n} g_i = 2n - 2. \tag{8.11}$$

Conversely, we find that this equality is also sufficient:

**Lemma 8.5.1** *Let $g_1, \ldots, g_n$ be a sequence of positive integers satisfying (8.11). Then there exists a tree on $n$ vertices with these predetermined degrees.*

*Proof.* Let $g_1, \ldots, g_{n+1}$ be a sequence with

$$\sum_{i=1}^{n+1} g_i = 2(n+1) - 2 = 2n. \tag{8.12}$$

Not all of the $g_i$ can be equal 1, since otherwise

$$\sum_{i=1}^{n+1} g_i = \sum_{i=1}^{n+1} 1 = n + 1 < 2n.$$

Not all of the $g_i$ can be greater than 1, since otherwise

$$2n = \sum_{i=1}^{n+1} g_i \geq \sum_{i=1}^{n+1} 2 = 2(n+1).$$

Hence, without loss of generality, we may assume that $g_{n+1} = 1$ and $g_n > 1$. Define $g'_1, \ldots, g'_n$ by

$$g'_i = g_i \tag{8.13}$$

for $i = 1, \ldots, n-1$, and

$$g'_n = g_n - 1. \tag{8.14}$$

For this sequence it holds

$$\begin{aligned}\sum_{i=1}^{n} g'_i &= \sum_{i=1}^{n-1} g_i + (g_n - 1) + (g_{n+1} - 1) \\ &= \sum_{i=1}^{n+1} g_i - 2 \\ &= 2n - 2.\end{aligned}$$

By the induction assumption there is a tree $T' = (V' = \{v_1, \ldots, v_n\}, E')$ such that $g(v_i) = g'_i$. Then the tree $T = (V' \cup \{v_{n+1}\}, E' \cup \{\underline{v_n v_{n+1}}\})$ fulfills the assertion.

□

Hence, the number of different trees increases exponentially, but not faster:

**Theorem 8.5.2** *Let $g_1, ..., g_n$ be a sequence of positive integers and denote by $t(n, g_1, ..., g_n)$ the number of different labelled trees $T = (\{v_1, ..., v_n\}, E)$ of $n$ vertices with the degree sequence*

$$g_T(v_i) = g_i \tag{8.15}$$

*for $i = 1, ..., n$. Then*

$$t(n, g_1, ..., g_n) = \frac{(n-2)!}{\prod_{i=1}^{n}(g_i - 1)!} \tag{8.16}$$

*if (8.11) holds, and*

$$t(n, g_1, ..., g_n) = 0 \tag{8.17}$$

*otherwise.*

*Proof.*([9], following the idea of 8.5.1)
In view of 8.5.1 we know that $t(n, g_1, ..., g_n) > 0$ if and only if (8.11) holds.
Without loss of generality, we may assume that

$$g_1 \geq g_2 \geq \ldots \geq g_n.$$

Then $v_n$ must be a leaf.
Let $C_i$ be the collection of all trees $T$ with vertices $v_1, \ldots, v_n$ and degrees $g_j = g_T(v_j)$, such that the leaf $v_n$ is adjacent to $v_i$. Assuming $g_i \geq 2$ we have

$$|C_i| = t(n-1, g_1, \ldots, g_{i-1}, g_i - 1, g_{i+1}, \ldots, g_{n-1}).$$

Since the collection of all trees is the union of the sets $C_i$ for $g_i \geq 2$ we obtain, by the addition principle,

$$t(n, g_1, ..., g_n) = \sum_{g_i \geq 2} t(n-1, g_1, \ldots, g_{i-1}, g_i - 1, g_{i+1}, \ldots, g_{n-1}). \tag{8.18}$$

Now, we use induction. The theorem is true for $n = 2$. Assume that $n \geq 3$ and that the theorem is true for $n - 1$. Then

$$\begin{aligned} & t(n, g_1, ..., g_n) \\ = \quad & \sum_{g_i \geq 2} t(n-1, g_1, \ldots, g_{i-1}, g_i - 1, g_{i+1}, \ldots, g_{n-1}) \\ = \quad & \sum_{g_i \geq 2} \frac{(n-3)!}{(g_1 - 1)! \cdots (g_{i-1} - 1)!(g_i - 2)!(g_{i+1} - 1)! \cdots (g_{n-1} - 1)!} \\ = \quad & \frac{(n-2)!}{(g_1 - 1)! \cdots (g_{n-1} - 1)!} \\ = \quad & \frac{(n-2)!}{(g_1 - 1)! \cdots (g_n - 1)!}, \end{aligned}$$

where we use 2.2.12.

□

Summing up over all degree sequences satisfying (8.11),

$$\begin{aligned} \# \text{ labelled trees} \quad &= \quad \sum_{(8.11)} t(n, g_1, ..., g_n) \\ &= \quad \sum_{(8.11)} \frac{(n-2)!}{\prod_{i=1}^{n} (g_i - 1)!} \\ &= \quad n^{n-2}, \end{aligned}$$

by 1.6.2 or 2.2.12, and we have one of the most beautiful formulas in enumerative combinatorics:

**Theorem 8.5.3** *(Cayley's tree formula, [21]) The number of different labelled trees with $n$ vertices equals $n^{n-2}$.*

### 8.5.2 The Prüfer code

Prüfer [66] established a bijection between trees and sequences of $n-2$ integers between 1 and $n$, providing a constructive proof of Cayley's result. This bijection can then be exploited to give algorithms for systematically generating labelled trees. More precisely: The strategy of the proof is to establish a one-to-one correspondence between the labelled tree and the Prüfer code, which is a sequence of length $n-2$ of integers between 1 and $n$, with repetitions allowed; in other words, a member of $\{1, \ldots, n\}^{n-2}$. Algorithmically this coding is described by

**Algorithm 8.5.4** *Let $T = (V = \{v_1, \ldots, v_n\}, E)$ be a labelled tree. Then the Prüfer code for $T$ can be constructed by performing the following steps:*

```
1. Initialize $T$ to be the given tree;
2. For $i = 1$ to $n - 2$ do
       Let $v$ be the leaf with the smallest label;
       Let $s_i$ be the label of the only neighbour of $v$;
       $T := T[V \setminus \{v\}]$;
3. The code is $(s_1, \ldots, s_{n-2})$.
```

We will now use the correspondence between Prüfer codes and labelled trees to generate trees. We first note that the following decoding procedure maps a given Prüfer code to a labelled tree:

**Algorithm 8.5.5** *A Prüfer code $P$ is given. Then a labelled tree $T = (V, E)$ can be constructed by performing the following steps:*

```
1. Initialize the list $P$ as the input;
2. Initialize the list $V$ as $1, \ldots, n$;
3. Initialize $T$ as the forest of isolated vertices on $V$;
4. For $i = 1$ to $n - 2$ do
       Let $k$ be the the smallest number in list $V$ that is not in list $P$;
       Let $j$ be the first number in list $P$;
       Add an edge joining the vertices labelled $k$ and $j$;
       Remove $k$ from list $V$;
       Remove the first occurrence of $j$ from list $P$;
```

5. *Add an edge joining the vertices labelled with the two remaining numbers in the list $V$.*

It is not hard to see that the decoding procedure 8.5.5 is the inverse of the encoding procedure 8.5.4. Altogether this establishes again 8.5.3.
Combining all these considerations gives the following:

**Algorithm 8.5.6** *Let $n$ be an integer with $n \geq 2$. Then the following algorithm generates all trees with $n$ labelled vertices:*

1. *Generate, by simple counting, all Prüfer codes in $\{1, \ldots, n\}^{n-2}$;*
2. *For each code apply 8.5.5.*

## 8.6 UNLABELLED TREES

Let $t(n)$ be the number of non-isomorphic (unlabelled) trees with $n$ vertices. By considering all $n!$ labellings, we have $n! \cdot t(n) \geq n^{n-2}$. Hence,

$$t(n) \geq \frac{n^{n-2}}{n!} \geq \frac{e^n}{en^3}. \tag{8.19}$$

On the other hand, we will describe a (nonoptimal) technique, involving drawing a tree in the plane: Let $n > 1$ be an integer. Remember that a tree code $w$ (with respect to $n-1$) is a sequence in $\mathbb{B}^{2(n-1)}$ with the following properties:

(i) In each prefix of $w$ the number of 1s is at least the number of 0s;
In particular, the first letter in $w$ must be 1;

(ii) The number of 1s in $w$ equals the number of 0s;
In particular, the last letter in $w$ must be 0.

**Algorithm 8.6.1** *Let $w$ be a tree code with respect to $n-1$. Then draw a tree by the following algorithm:*

1. *Put a vertex as the origin;*

2. *Read w letter by letter and*
   *if you see a 1 then draw a new edge to a new vertex;*
   *if you see a 0 then move back by one edge toward the origin.*

Thus the tree is described by its tree code. Hence, after generating all tree codes, we can generate all unlabelled trees with $n$ vertices.
Note that the tree code is far from optimal; every unlabelled tree has many different codes. For instance all the codes 11010010, 10110100, 11101000, 10101100, 11011000 and 11100100 generate the same tree.

According to a difficult result of Pólya, compare [43], the number of unlabelled trees is asymptotic completely determined:

$$t(n) \approx \frac{c \cdot a^n}{n^{5/2}}, \tag{8.20}$$

where $a = 2.9557\ldots$ and $c = 0.5349\ldots$.

Otter [62] finds the following values for the number of trees.

| Number $n$ of vertices | Number $t(n)$ of trees |
|---:|---:|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 2 |
| 5 | 3 |
| 6 | 6 |
| 7 | 11 |
| 8 | 23 |
| 9 | 47 |
| 10 | 108 |
| $\vdots$ | |
| 24 | 39299897 |

For these numbers and other facts about counting of trees see [45].

## 8.7 BINARY TREES

A tree in which each vertex has degree one or three is called a binary tree. Binary trees play an important role in the theory of evolution, since it is assumed that a phylogenetic tree is a "bifurcation" tree. This follows from the assumption that evolution is driven by bifurcation events.[6]

**Observation 8.7.1** *A binary tree for $n$ leaves has an even number of vertices, namely $2n-2$.*

With this in mind we have the following consequences of 8.5.2, showing that the number of possible phylogenetic trees increases rather dramatically as the number of taxa increases.

**Theorem 8.7.2 (a)** *The number of binary trees with $n$ labelled leaves and $n-2$ labelled internal vertices is*

$$\frac{(2n-4)!}{2^{n-2}}. \tag{8.21}$$

**(b)** *(Cavalli-Sforza, Edwards [19]) The number of binary trees with $n$ labelled leaves and $n-2$ unlabelled internal vertices is*

$$(2n-5)!! = 1 \cdot 3 \cdot 5 \cdots (2n-5) = \Omega\left(\left(\frac{2n}{3}\right)^{n-2}\right). \tag{8.22}$$

The following table, the result is applied to phylogenetic trees:

| Number of taxa | Number of binary trees | Comment |
|---:|---:|---|
| 4 | 3 | |
| 10 | 10395 | |
| 22 | $\approx 3 \cdot 10^{23}$ | Almost a mole of trees |
| 50 | $\approx 3 \cdot 10^{74}$ | More trees than the number of atoms in the universe |
| 100 | $\approx 2 \cdot 10^{182}$ | out of any range |

[6] In practice phylogenetic trees are allowed to be multifurcating when the bifurcations are sufficiently close together.

A helpful description of binary trees with labelled leaves is given by the following procedure: Let $T = (V, E)$ be a binary $N$-tree for $N = \{v_1, \ldots, v_n\}$.

1. If $n = 2$, then write $T$ as $(v_1 v_2)$; otherwise,
2. Let $v_i$ and $v_j$ be two leaves of $T$ which are adjacent to the same vertex $v$. Then
   (i) Delete the leaves $v_i$ and $v_j$, and its incident edges;
   (ii) Replace the vertex $v$ by $(v_i v_j)$, which is now a leaf;
   (iii) Consider the new tree with $n - 1$ leaves and repeat the procedure.

Clearly, this procedure gives a simple written description of the tree, called the "bracket" or Newick format. But note that it is not unique, for example for the one $N$-tree for $n = 3$ we have the descriptions $((v_1 v_2) v_3)$ and $(v_1 (v_2 v_3))$ and $((v_1 v_3) v_2)$.

## 8.8 ROOTED TREES

The most important point in a phylogenetic tree is its root. In a rooted tree exactly one distinguished vertex is marked as the root.

For each of the labelled trees we have $n$ rooted trees, because any of the $n$ vertices can be made a root. Hence, as a consequence of Cayley's tree formula we find:

**Corollary 8.8.1** *The number of different rooted labelled trees with n vertices equals $n^{n-1}$.*

A unique path leads from the root to any other vertex of the tree. Let $w$ be the root and $v$ be an arbitrary vertex in a rooted tree $T = (V, E)$. The length of the path[7] from $w$ to $v$ is called the level of $v$:

$$\text{level}(v) = \rho(w, v). \tag{8.23}$$

The depth of the tree itself is defined by

$$\text{depth}(T) = \max\{\text{level}(v) : v \in V\}. \tag{8.24}$$

[7] Remember that in this case the length is the number of edges in the path.

If $T$ is a rooted tree, then it is customary to draw $T$ with root $w$ at the top, at level 0. The vertices adjacent to $w$ are placed one level 1. Any vertex adjacent to a vertex of level 1 is at level 2, and so on. In general, every vertex at level $i > 0$ is adjacent to exactly one vertex at level $i - 1$.

We may consider a rooted tree $T = (V, E)$ as a digraph if we direct the edges $\underline{vv'} \in E$ from $v$ to $v'$ if and only if $\text{level}(v') = \text{level}(v) + 1$, where $w$ is the root of $T$. Then $g^{in}(w) = 0$ characterizes the root, and $g^{out}(v) = 0$ characterizes the leaves of $T$.
In this sense we have an ancestor/successor-relation for the vertices of a rooted tree. In particular, the root is the common ancestor of all vertices of the tree. In other words, a rooted tree has a vertex identified as the root from which ultimately all other vertices descend.

For a rooted tree $T = (V, E)$ a natural partial order $\leq_T$ on the set $V$ of vertices is obtained by setting $v \leq_T v'$ if
- The path from the root of $T$ to $v'$ includes $v$ or, equivalently,
- $v'$ is the successor of $v$, and $v$ is the ancestor of $v'$.

Obviously, $v \leq_T v'$ implies $\text{level}(v) \leq \text{level}(v')$ (but not vice versa).
Let $T = (V, E)$ be a rooted $N$-tree and let $N'$ be a subset of $N$. We will refer to the unique vertex $v$ of $T$ that is the greatest lower bound of $N'$ under the order $\leq_T$ as the last universal common ancestor of $N'$ in $T$. That means
- $v$ is an ancestor for each vertex in $N'$, and
- $\text{level}(v) = \max\{\text{level}(v') : v' \text{ ancestor for each member in } N'\}$.

A tree $T$ is called a rooted binary tree if for its vertices

$$g_T(v) = \begin{cases} 1 & : \text{ if } v \text{ is a leaf} \\ 2 & : \text{ if } v \text{ is the root} \\ 3 & : \text{ otherwise} \end{cases}$$

holds. In other words, we create a rooted binary tree from a binary tree by choosing an edge and place the root there. This procedure is called rooting a tree.

Rooted trees are representations for evolutionary relationships. For a rooted $N$-tree $T$ we view the edges as being directed away from the root, and then regard $T$ as describing the evolution of the set $N$ of given "names" from a common (hypothetical) ancestral name; the other internal vertices of $T$ correspond to

further ancestral names.[8,9]
Hence, the most important point in a phylogenetic tree is its root. The root is placed at this position to indicate that,

(i) it corresponds to the (theoretical) last universal common ancestor of everything in the tree;

(ii) gives directionality to evolution within the tree; and

(iii) it identifies which groups of vertices are "true", given if the root does not lie within a group.

The question is: On which edge should the root be placed? There are three popular ways to find this position:

1. On the longest edge[10].
2. In the middle of the longest path between two leaves.
3. An "outgroup" can be added to the set of given points. Then the root is placed at the bifurcation between the outgroup and the main group.

Next, we will discuss the relationship between the number of leaves in a rooted binary tree and its depth. It is not hard to see that

**Observation 8.8.2** *Let $T$ be a rooted binary tree of depth $d$. Then $T$ has at least $d + 1$ and at most $2^d$ leaves.*

Conversely, the depth of such a tree with $n$ leaves lies approximately between $\Omega(\log n)$ and $O(n)$.

Now we count rooted trees. Together with 8.7.2 we have

[8] Unrooted phylogenetic trees are also biologically relevant since they are typically what tree reconstruction methods generate.

[9] Rooting a tree has a strong relationship to the molecular clock; but especially, proteins evolve at different rates, making it difficult to relate the (evolutionary) distance to the historical time.

[10] This approach of course requires that there is a length-function for the graph.

**Theorem 8.8.3** *The number of rooted binary trees with $n$ labelled leaves and unlabelled internal vertices (i.e. rooted binary $N$-trees having $|N| = n$) is*

$$(2n-3)!! = 1 \cdot 3 \cdot 5 \cdots (2n-3) = \Omega\left(\left(\frac{2n}{3}\right)^{n-1}\right). \tag{8.25}$$

# 9

# TRAVELLING ROUND A GRAPH

## 9.1 EULERIAN CYCLES

If we allow more than one edge in $E$ to join two vertices in $V$, meaning that there are parallel edges in the graph, we shall call the pair $(V, E)$ a multigraph. In this sense, any graph is a also a multigraph.

Let $G$ be a graph. A Eulerian chain of $G$ (Eulerian cycle of $G$, respectively) is defined as a chain (cycle, respectively) that uses each edge of $G$ exactly once.[1]
A graph which contains a Eulerian cycle is called a Eulerian graph.
One of the oldest combinatorial problems, accredited to Euler and written in the teminology of graph theory, can be stated as follows: When does a multigraph have a Eulerian chain or a Eulerian cycle? The answer is:

**Theorem 9.1.1** *(Euler) A multigraph has an Eulerian cycle if and only if it is connected and all vertices have even degree.*

*Proof.* Suppose that the graph $G$ is Eulerian with the Eulerian cycle $C$. Then, of course, $G$ must be connected; and the chain $C$ enters and leaves any vertex the same number of times without repeating any edge, hence with an even degree.

The converse of 9.1.1 is given by an algorithm for finding such a cycle effectively: Start with a cycle through the multigraph and add a ”detour” cycle until all edges are in the tour.

[1] Note that an Eulerian cycle is not a cycle in the usual sense, since it can contain a vertex more than once.

□

**Algorithm 9.1.2** *(Hierholzer, compare [48], [54]) Let $G = (V, E)$ be a Eulerian (multi-) graph. Choose a vertex $v_1$ arbitrarily and apply the following recursive procedure Euler$(G, v_1)$ to find a Eulerian cycle:*

```
1. Set $C := v_1$; $v := v_1$;
2. If $g_G(v) = 0$ then goto 4.
   else let $w \in N_G(v)$, choose one edge $e = \underline{vw}$;
3. Set $C := C, e, w$ and $v := w$;
   Set $E := E \setminus \{e\}$;
   goto 2.;
4. Let $C = v_1, e_1, v_2, e_2, \dots, v_k, e_k, v_{k+1}$;
   For $i := 1$ to $k$ do $C_i :=$ Euler$(G, v_i)$;
5. Set $C = C_1, e_1, C_2, e_2, \dots, C_k, e_k, v_{k+1}$.
```

Theorem 9.1.1 also has several consequences:

(i) A multigraph has an (open) Eulerian chain if and only if it is connected and has exactly two vertices of odd degree.

(ii) Any connected graph contains a chain that uses each edge exactly twice.

(iii) A (multi-)digraph $G$ has an Eulerian cycle if and only if $G$ is connected and for every vertex the indegree equals the outdegree.

For applications of Eulerian graphs in breaking polynecleotides compare [67].

## 9.2 HAMILTONIAN GRAPHS AND $K$-ORDERS

A question similar to the problem of Euler was raised by Hamilton. Let $G$ be a graph. A Hamiltonian cycle is a cycle that contains all vertices of $G$. The problem is to decide whether or not $G$ has a Hamilton cycle; if so then $G$ is called a Hamiltonian graph.

Hamilton's problem sounds quite similar to Euler's, but this is not the case, as there is an essential difference:

- An Eulerian cycle traverses every edge exactly once, but may repeat vertices, while Hamiltonian cycle visits each vertex exactly once.
- An Eulerian cycle contains also all vertices of the graph, but a Hamiltonian cycle need not contain all edges.

Although it is clear that only connected graphs can be Hamiltonian, there is no simple criterion to tell us whether or not a graph is Hamiltonian as there is for Eulerian graphs. And indeed, no efficient algorithmic method is known to check whether a given graph has a Hamiltonian cycle. Karp [50] showed that the problem whether or not a graph is Hamiltonian is very difficult to solve in the sense of computational complexity.

**Theorem 9.2.1** *$Q_n$ is Hamiltonian.*

The proof uses induction, compare [36].

There is an important application of 9.2.1 in coding theory. A Gray code is a cyclic arrangement of binary sequences such that any pair of adjacent sequences differ in only one position. Example: 000 → 010 → 110 → 100 → 101 → 111 → 011 → 001 →. This sequence corresponds to a Hamilton cycle in $Q_3$.

A weaker question as Hamilton's is the following. Let $G = (V, E)$ be a connected graph with $n$ vertices. $\rho(v, v')$ denotes the length of a shortest path between the vertices $v$ and $v'$.[2]
Let $k$ be a positive integer. A $k$-order for $G$ is a permutation $\pi$ of $\{1, \ldots, n\}$ such that

$$\rho(v_{\pi(i)}, v_{\pi(i+1)}) \leq k \tag{9.1}$$

for $i = 1, \ldots, n-1$, and

$$\rho(v_{\pi(n)}, v_{\pi(1)}) \leq k. \tag{9.2}$$

A 1-order is a Hamilton cycle.
We will prove the surprising result that each connected graph has a 3-order, and start with a stronger statement.

[2] This function is defined by the metric closure of $G$ with the length-function $f \equiv 1$.

**Lemma 9.2.2** *(Karaganis) Let $T = (V, E)$ be a tree with $n$ vertices, and let $v, v'$ be two vertices. Then there is a order $v = v_1, v_2, \ldots, v_{n-1}, v_n = v'$ such that*

$$\rho(v_i, v_{i+1}) \leq 3 \tag{9.3}$$

*for $i = 1, \ldots, n-1$.*

*Proof.* We use induction over $n$.
The lemma is true for $n = 2, 3$. Now we assume that it is true for all trees with less than $n$ vertices.
Let $v = v_1, v_2, \ldots, v_{r-1}, v_r = v'$ be the path interconnecting $v$ with $v'$.

$$\begin{aligned} G_1 &= G - \underline{v_1 v_2}, \\ G_i &= G - \underline{v_{i-1} v_i} - \underline{v_i v_{i+1}}, \text{ for } i = 2, \ldots, r-1 \\ G_r &= G - \underline{v_{r-1} v_r} \end{aligned}$$

is a forest, where the tree $G_i$ contains the vertex $v_i$. In view of the induction hypothesis for each $i = 1, \ldots, r$ there is a order $v_i = v_1^i, v_2^i, \ldots, v_{n_i}^i$ in $G_i$ such that

$$\rho(v_j^i, v_{j+1}^i) \leq 3 \text{ for } j = 1, \ldots, n_i - 1 \tag{9.4}$$

$$\rho(v_i, v_{n_i}^i) = 1, \tag{9.5}$$

where $n_i$ denotes the number of vertices in $G_i$.
Then we construct the desired order by

$$\begin{aligned} &v = v_1 = v_1^1, \ldots, v_{n_1}^1, v_2 = v_1^2, \ldots, v_{n_2}^2, \ldots, v_{r-1} = v_1^{r-1}, \ldots, v_{n_{r-1}}^{r-1}, \\ &v_{n_r}^r, v_{n_r-1}^r, \ldots, v_1^r = v_r = v'. \end{aligned} \tag{9.6}$$

□

**Theorem 9.2.3** *(Sekanina) Each connected graph has a 3-order.*

*Proof.* First use a spanning tree of the graph, then apply 9.2.2 for two adjacent vertices.

□

## 9.3 THE SHORTEST SUPERSTRING PROBLEM

Remember the longest common subsequence problem, which we solved by a dynamic programming approach. The converse of this problem is

**The problem of shortest common supersequence**
**Given:** A set of sequences over the same alphabet.
**Find:** A shortest sequence that contains each of the given sequences as a subsequence.

This (in short:SCS) problem plays a favourite role in DNA sequencing. In fact, DNA sequencing is routinely done by sequencing large numbers of relatively short fragments and then finding a short common supersequence.

Let $S = \{w_1, \ldots, w_n\} \subseteq A^\star$ be a set of strings (words) over the alphabet $A$.

Throughout the discussion of superstrings, we assume that no string in $S$ is a substring of any other string in $S$. Any such substring can be efficiently detected (How?) and consequently removed. After that the problem has the same solution as before, which means that a SCS for the remaining strings is also a SCS of the original set.

For two strings $w, w' \in A^\star$ we define the string

$$\text{Merge}(w, w') = xyz, \tag{9.7}$$

where

(i) $y$ is a suffix of $w$;

(ii) $y$ is a prefix of $w'$; and

(iii) $|y|$ is maximal.

$y$ is called the overlap of $w$ and $w'$, and written by $y = \text{overlap}(w, w')$.
$\text{prefix}(w, w')$ is the prefix of $w$ ending at the start of the overlap. Of course,

$$|\text{prefix}(w, w')| = |w| - |\text{overlap}(w, w')| \tag{9.8}$$

**Observation 9.3.1** *The function |prefix(., .)| satisfies the triangle inequality.*[3]

Let $\pi$ be a permutation of $\{1, \ldots, n\}$, then consider the following string:

$$w[\pi] = \text{prefix}(w_{\pi(1)}, w_{\pi(2)})\text{prefix}(w_{\pi(2)}, w_{\pi(3)}) \ldots \text{prefix}(w_{\pi(n-1)}, w_{\pi(n)})w_{\pi(n)}. \tag{9.9}$$

$w[\pi]$ is the concatenation of the nonoverlapping prefixes of the pairs of adjacent strings, followed by the full string of the last index.

**Theorem 9.3.2** *For a set $S = \{w_1, \ldots, w_n\} \subseteq A^\star$ of strings and the permutation $\pi$ the string $w[\pi]$ is a superstring of $S$ with length*

$$|w[\pi]| = \sum_{i=1}^{n-1} |\mathit{prefix}(w_{\pi(i-1)}, w_{\pi(i)})| + |w_{\pi(n)}|. \tag{9.10}$$

Consequently, looking for a SCS is the search of a permutation $\pi$ such that $|w[\pi]|| = \min$.

For $S$ we define the distance-graph $G = (V, E)$ by the following definitions:

(a) $V = S$;

(b) $E = V \times V$, that means, really $G$ is a digraph including loops;

(c) There is a length-function $c : E \to I\!N$ with

$$c(w, w') = \begin{cases} |\text{prefix}(w, w')| & : \quad w \neq w' \\ |w| & : \quad w = w' \end{cases}$$

Then, looking for a SCS is the search of minimal travel through the distance graph.

The shortest superstring problem is very hard in the sense of computational complexity, compare [84], and the approach above need a great amount of time. In view of this fact we are interested in an approximation strategy for the shortest superstring problem.

For $S$ we define the overlap-graph $G = (V, E)$ by the following definitions:

[3] But is not a metric, since the symmetry fails.

(a) $V = S$;

(b) $E = V \times V \setminus \{(w,w) : w \in V\}$, that means, really $G$ is a digraph;

(c) There is a length-function $c : E \to I\!N$ with

$$c(w,w') = |\text{overlap}(w,w')|$$

for $w \neq w'$.

In view of (9.8) we can the overlap-graph derive from the distance-graph.

**Algorithm 9.3.3** *Let $S = \{w_1, \ldots, w_n\} \subseteq A^\star$ be a set of strings. Then a superstring can be found by the following strategy.*
*While $|S| > 1$ do*

```
1. Find $w_i, w_j \in S$ such that
       $|overlap(w_i, w_j)| = \max\{|overlap(w, w')| : w, w' \in S\}$
   for $i \neq j$;
2. $w = Merge(w_i, w_j)$;
3. $S := S \setminus \{w_i, w_j\} \cup \{w\}$.
```

*The remaining string is the searched superstring.*

In other terms, we sequentially choose the longest edge that does not form a cycle with already chosen edges.

As example consider the three-string set

$$S = \{c(ab)^m, (ba)^m, (ab)^m c\} \subseteq \{a,b,c\}^\star,$$

with a given positive integer $m$.
After the first run of 9.3.3 we found the two-string set

$$S^{\text{new}} = \{c(ab)^m c, (ba)^m\}.$$

Consequently, the output of 9.3.3 is the string

$$c(ab)^m c(ba)^m \text{ or } (ba)^m c(ab)^m c$$

each of length $4m+2$. On the other hand, a SCS for the original set is (obviously)

$$ca(ba)^m bc$$

of length $2m+4$.

# 10

# CLASSIFICATIONS

Naming is classifying.

Brian Everitt

In the widest sense, a classification scheme may represent simply a convenient method for organizing a large set of data so that the retrieval of information may be made more efficiently. In this sense, classification is the begin of all science.

One of the most basic abilities of living creatures is the grouping of similar objects to produce a classification. This has been a preoccupation since the very first biological investigations. The classification of animals and plants has clearly played an important role in the fields of biology as a basis for Darwin's theory of evolution. The theory and practice of classifying organisms is generally known as taxonomy. In 1737 Linnaeus published his work *Genera Plantarum*; he wrote:

> All the real knowledge which we possess, depends on methods by which we distinguish the similar from the dissimilar. The greater number of natural distinctions this method comprehends the clearer becomes our idea of things. The more numerous the objects which employ our attention the more difficult it becomes to form such a method and the more necessary.
> ...
> For we must not join in the same genus the horse and the swine, though both species had been one hoof'd nor separate in different genera the

> goat, the reindeer and the elk, tho' they differ in the form of their horns. We ought therefore by attentive and diligent observation to determine the limits of the genera, since they cannot be determined a priori. This is the great work, the important labour, for should the genera be confused, all would be confusion.

In other words, taxonomy is necessary, but must be done carefully.[1] A rough classification of the world of organisms is

$$\begin{aligned} \text{organisms} &= \{\{\text{prokaryota}\} = \{\text{archea}\} \cup \{\text{bacteria}\}, \\ &\quad \{\text{eukaryota}\} = \{\text{protista}\} \cup \{\text{plantae}\} \cup \{\text{fungi}\} \cup \\ &\quad \{\text{animalia}\}\}. \end{aligned}$$

Classification has played a central role in other fields too. In particular, the classification of the elements in the periodic table, given by Mendeleyev 150 years ago, has had a profund impact on the understanding of the structure of atoms. More examples are given in [31].[2]
As example in linguistics we give a very partial representation of branches of Indo-European language family.

| | | |
|---|---|---|
| Indo-European | Germanic | German |
| | | English |
| | | Danish |
| | Slavic | Russian |
| | | Polish |
| | Indo-Iranian | Persian |
| | | Hindi |

For a classification of languages see [23].

**The Problem of Classification**

**Given:** A collection of objects, each of which is described by a set of characters or variables.

**Derive:** A useful (whatever that means) division into a hierarchy of classes.

[1] Linnaeus' purpose was not evolutionary, but rather to provide a set of universal names. However it turned out that the hierarchical nature of his system has considerable similarity with the modern phylogenetic view.

[2] To radically simplify, in the cases, human beings and behaviour may be classified into classes named by *low*, *medium* and *high*.

Numerical techniques for devising classification must be objective and stable, which means

- Objective in the sense that the analysis of the same set of objects (individuals) by the same family of methods will produce the same classification.
- Stable in the sense that classifications remains the same under a wide varity of additions of individuals.

## 10.1 CLASSIFICATION AND EVOLUTION

Evolution implies that many different species have a common ancestor and that all forms of life probably stem from the same remote beginnings. Once these relationships are understood, they are summarized by grouping species into collections of related organisms. We will describe the structures underlying these relationships.

Let $N$ be the set of of extant species (genes) and let $N^+$ be the set of all past and present species (genes). Then we consider the binary operation $\star : N^+ \times N^+ \to N^+$ defined by

$$v \star v' = \text{most recent common ancestor of } v \text{ and } v', \tag{10.1}$$

for $v, v' \in N^+$, whereby $v \star v = v$.
Moreover, define for $v \in N^+$ the set $N(v)$ as the set of extant species (genes) descented from $v$. Then, we have to assume that for two species (genes) $v$ and $v'$, either the two sets $N(v)$ and $N(v')$ are disjoint or one is contained in the other.

**Observation 10.1.1** *Defining the sets as above, the conditions*

*(i)* $N(v) \cap N(v') \neq \emptyset$*, and*

*(ii)* $N(v) \subseteq N(v')$ *or* $N(v') \subseteq N(v)$*,*

*are equivalent.*

The theory of evolution is concerned with the extraordinary diversity of life on Earth. The diversity of the living world is staggering: more than 2 million

existing species of plants and animals have been named and described; and many more remain to be discovered - until up to 10 times this number according to some estimates. What is impressive is not just the numbers but also the incredible heterogeneity. These virtually infinite variations of life are the fruit of the evolutionary process.
Taxonomy is the classification of organisms for the first aspect in any view of the life. The classification of animals and plants played an important role as a basis for Darwin's theory of evolution.
Taxonomy is necessary to describe the diversity of living organisms, whereby the diversity of genomes is twofold:

- The presence of numerous species on Earth; and
- The polymorphism within each species.

There are many reasons why knowledge of the biodiversity is necessary, compare [37], and [57].[3] There are several subquestions:

(i) How many species are there?

(ii) How many go extinct? In both the past and in the present. How many are lost every year?

(iii) How long did species typically survive?

(iv) How many are newly created?

(v) How much of evolutionary history is knowable?

For using evolutionary history for describing the biodiversity see [73].

## 10.2 CLASSIFICATIONS AND TREES

A classification is the formal naming of a group of individuals. In the sense of set theory a classification $\mathcal{C}$ of a (finite) set $N$ of individuals is given by a collection of subsets of $N$ satisfying

(i) $\emptyset \notin \mathcal{C}$;

[3] In particular, there no successful vaccine to prevent or halt HIV infection. In part, this is because of the high genetic diversity of HIV, see [29].

(ii) $N \in \mathcal{C}$;

(iii) $\{v\} \in \mathcal{C}$ for any $v \in N$; and

(iv) For any two members $N'$ and $N$" of $\mathcal{C}$ it holds that

$$N' \cap N" \in \{N', N", \emptyset\}. \tag{10.2}$$

In other words, any two sets in $\mathcal{C}$ are disjoint or one is contained in the other (see 10.1.1).

A member of a classification is called a class or a cluster of $N$.

Let $T$ be an $N$-tree rooted by the vertex $w$. Then we create a collection $\mathcal{C}$ of classes for the set $N$ in the following way:

1. For each leaf $v$ of $T$ put $\{v\}$ in $\mathcal{C}$;
   Label the vertex $v$;
2. Let $v \neq w$ be an unlabelled vertex adjacent to exactly one other unlabelled vertex. All other neighbors $v_1, \ldots, v_k$ of $v$ are labelled and belong to classes $N_1, \ldots, N_k$ in $\mathcal{C}$, respectively. Then
   - Put $\bigcup_{i=1}^{k} N_i$ in $\mathcal{C}$, and
   - Label $v$;
3. Label $w$;
   Put $N$ in $\mathcal{C}$.

Conversely, if we have a collection $\mathcal{C}$ of classes of the set $N$ with the properties that $\{v\} \in \mathcal{C}$ for each element $v \in N$ and $N \in \mathcal{C}$, we can form a tree $T$ by:

1. Each class of $\mathcal{C}$ is a vertex of $T$;
2. Two vertices $N_1$ and $N_2$ are adjacent if and only if
   - $N_1 \cap N_2 \in \{N_1, N_2\}$, and
   - there is no class $N'$ such that $N_j \cap N' \in \{N_j, N'\}$ for $j = 1, 2$.
   (That means, $N_1$ must be the maximal proper subset of $N_2$ or vice versa.)

Summing up all these observations, we have the following fundamental equivalence between classifications and rooted trees.

**Observation 10.2.1** *There is a one-to-one correspondence between the collection of classifications for a set $N$ and the collection of rooted $N$-trees.*

In other words, classifications for a set $N$ and rooted $N$-trees contain essentially the same information.[4] The following proposition describes the equivalence between a classification and a desired collection of splits. The proof is an application of 10.2.1.
A pair $\{N_1, N_2\}$ and $\{M_1, M_2\}$ of splits for $N$ is called compatible if at least one of the sets $N_1 \cap M_1$, $N_1 \cap M_2$, $N_2 \cap M_1$ and $N_2 \cap M_2$ is the empty set.

**Observation 10.2.2** *(Semple and Steel [75]) Let $N$ be finite set. $\mathcal{C}$ is a classification for $N$ if and only if the collection*

$$\mathcal{S} = \{\{N_1, N_2\} : N_1 \in \mathcal{C} \setminus \{N\}, N_2 = N \setminus N_1\} \tag{10.3}$$

*is a set of pairwise compatible splits on $N$; and vice versa.*

For instance, consider the set $N = \{a, b, c, d, e\}$. Coming from the (binary) $N$-tree $(((ab)c)(de))$ we have the split system

$$\begin{aligned} \mathcal{S} \quad = \quad & \{\{a, bcde\}, \{b, acde\}, \{c, abde\}, \{d, abce\}, \{e, abcd\}, \\ & \{ab, cde\}, \{abc, de\}\}. \end{aligned} \tag{10.4}$$

Using each of the three internal vertices as a root gives the following classifications:

$$\mathcal{C}_1 \quad = \quad \{a, b, c, d, e, ed, ced\} \cup \{N\} \tag{10.5}$$
$$\mathcal{C}_2 \quad = \quad \{a, b, c, d, e, ab, ed\} \cup \{N\} \tag{10.6}$$
$$\mathcal{C}_3 \quad = \quad \{a, b, c, d, e, ab, abc\} \cup \{N\}. \tag{10.7}$$

With 10.2.1 in mind, we have several considerations.

Firstly we determine the maximal number of sets in a classification. Let $T = (V, E)$ be a rooted $N$-tree with $|N| = n$, that $k$ internal vertices each of degree greater than 2, and a root $w$. Then 8.3.2 says that $k \leq n - 2$. Consequently,

[4] In view of this observation, each evolutionary tree implies a classification of the given names, but of course not vice versa. We saw in 8.8.2 that such a classification is not applicable in practice, since the depth of the tree lies between $\Omega(\log n)$ and $O(n)$ for $n = |N|$, and is obviously too big. Taxonomists are interested in trees with a constant depth. In particular Linnaeus' system has depth 8. Hence, in such systems the trees are not binary.

**Observation 10.2.3** *Let $\mathcal{C}$ be a classification for a set with $n$ elements. Then,*

$$n+1 \leq |\mathcal{C}| \leq 2n-1. \tag{10.8}$$

Secondly, we find a metric for rooted trees. This measure $\rho_C$ can be calculated easily, and the fact that it counts the different classes in the corresponding classifications is an indication of its biological relevance; see [45] and [68].
Recall that a rooted tree $T$ can be directed so that each edge is directed away from the root. Then for each edge $e$ of $T=(V,E)$ let $C(e)$ be the set of the marks of the vertices below $e$ in the tree. $C(e)$ is called the content of $e$, and

$$\mathcal{C}(T) = \{C(e) : e \in E\} \cup \{N\}. \tag{10.9}$$

In view of 1.7.1 we have

**Observation 10.2.4**

$$\rho_C(T_1,T_2) = |\mathcal{C}(T_1) \triangle \mathcal{C}(T_2)| \tag{10.10}$$

*is a distance between the two rooted $N$-trees $T_1$ and $T_2$.*

Third, a rooted $N$-tree $T$ generates a hierarchy: Let $d$ be the depth, and let $k$ be an integer between 0 and $d$. For any two leaves $v$ and $v'$ of $T$ we define the relation $v \sim_k v'$ if there is a path from $v$ to $v'$ in $T$ containing only vertices of a level $k$ or higher. It is easy to see that $\sim_k$ is an equivalence relation for any number $k$. $\mathcal{N}(k)$ denotes the family of the equivalence classes. Then we have a series $\mathcal{N}(0), \mathcal{N}(1), \ldots, \mathcal{N}(d)$ of partitions of $N$ with

(i) $\{N\} = \mathcal{N}(0)$ and $\mathcal{N}(d) = \{\{v\} : v \in N\}$.

(ii) For $k = 0, \ldots, d-1$ the class $\mathcal{N}(k+1)$ is finer than $\mathcal{N}(k)$.[5]

The first set $\mathcal{N}(0)$ consists of a group of ancestors, the last $\mathcal{N}(d)$ consists of individual leaves. Overall, we separate the "individuals" of $N$ into successively finer groupings.[6]

**Observation 10.2.5** *The following statements are pairwise equivalent.*

[5] The inclusion is strict, since all internal vertices of $T$ are of degree at least three.

[6] A nice illustration of this point of view is given in [38]:

- *$\mathcal{C}$ is a classification for $N$.*
- *$\mathcal{C}$ represents a rooted $N$-tree.*
- *$\mathcal{C}$ consists of a series of partitions for $N$ which become finer and finer.*

In this sense a classification is a hierarchy of partitions.

## 10.3 PAIR GROUPING

To create a classification of individuals from similarity values algorithmic approaches are popular. We create a proceeding of successive fusions of $n = |N|$ individuals into groups. These methods are well-known in cluster analysis. The related rooted trees are usually called dendrograms, compare [31]. The general idea of the algorithm is to repeatedly merge pairs of sets,[7] and so the technique is called a pair group method (PGM).

**Algorithm 10.3.1** *Let $N = \{v_1, \ldots, v_n\}$ be a set of individuals.*
*Firstly create a family $N_1 = \{v_1\}, \ldots, N_n = \{v_n\}$ of sets each containing a single element. Then creates a classification $\mathcal{C}$ by doing the following steps*

```
1. $N_1 = \{v_1\}, \ldots, N_n = \{v_n\} \in \mathcal{C}, \mathcal{R}$;
2. Find the nearest pair of distinct sets in $\mathcal{R}$, say $N_i$ and $N_j$;
3. Merge $N_i$ and $N_j$ to form $N'$;
   $\mathcal{C} := \mathcal{C} \cup \{N'\}$;
   $\mathcal{R} := \mathcal{R} \cup \{N'\} \setminus \{N_i, N_j\}$;
   Compute a new distance, or similarity from $N'$ to each of the other sets
   in $\mathcal{R}$;
```

| Biological | Postal |
|---|---|
| Domain | Old/New World |
| Kingdom | Country |
| Phylum | State/Province |
| Class | City |
| Order | Street |
| Family | Number |
| Genus | Last name |
| Species | First name |

[7] Only pairs are considerd in view of the bifurcation assumption of evolutionary processes.

*4. If $|\mathcal{R}| = 1$ then STOP, else go to 1.*

Obviously, this is a very general approach, and we have to specify several facts more precisely: What does "nearest" mean? How we can compute the new distance?

We start with the distance matrix $D = D(N, \rho) = (d_{ij})_{i,j=1,\ldots,n}$ for $N = \{v_1, \ldots, v_n\}$.[8] After each step of the procedure we compute a new matrix whose entries are inter-point and -class distances; in other words we convert the distance to a cluster-distance, that is a function $d : \mathcal{N} \times \mathcal{N} \to I\!R_{\geq 0}$, where $\mathcal{N}$ denotes a classification of $N$. The specifics of these computations distinguish the methods.
Note that we start with a distance matrix with $n(n-1)/2$ parameters. Since a tree is defined by $n-1$ parameters, we cut the number of parameters by the factor $n/2$. Thus we may lose some information.

The following is an input for which the algorithm fails: Let $N = \{v_1, \ldots, v_4\}$ and let $((v_1 v_2)(v_3 v_4))$ be a $N$-tree with the length-function equal to 3 for the edges adjacent to $v_1$ and $v_4$ and equal to 1 otherwise. Then the distance matrix is given by:

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ |
|---|---|---|---|---|
| $v_1$ | 0 | 4 | 5 | 7 |
| $v_2$ | | 0 | 3 | 5 |
| $v_3$ | | | 0 | 4 |
| $v_4$ | | | | 0 |

Algorithm 10.3.1 amalgamates $v_2$ and $v_3$, in violation of the fact that these vertices are separated in the original tree.

Our basic strategy will be linking the least distant pairs of taxa, say $N_i$ and $N_j$; followed by successively more distant taxa or classes of taxa. When two taxa are linked, they lose their individual identities and are subsequently referred to as a single class: $N' = N_i \cup N_j$; $\mathcal{R} := \mathcal{R} \setminus \{N_i, N_j\} \cup \{N'\}$. The process is complete when the last two classes are merged into a single class containing all of the original taxa. There are several PGM distinguished by the kind of

[8] Maybe derived from a metric $\rho$:

$$d_{ij} = \rho(v_i, v_j). \tag{10.11}$$

computation of the new distance matrix. More exactly, a new distance function (and matrix) is found by recalculating with $N'$ replacing $N_i$ and $N_j$ as follows: For all sets $K \in \mathcal{R} \setminus \{N'\}$ define

- Simple joining:
$$d(K, N') := \frac{d(K, N_i) + d(K, N_j) - d(N_i, N_j)}{2}; \tag{10.12}$$
- UPGMA (unweighted pair group method with arithmetic mean):
$$d(K, N') := \frac{d(K, N_i) + d(K, N_j)}{2}. \tag{10.13}$$
- WPGMA (weighted pair group method with arithmetic mean):
$$d(K, N') := \frac{|N_i| \cdot d(K, N_i) + |N_j| \cdot d(K, N_j)}{|N_i| + |N_j|}. \tag{10.14}$$

  Note that if the sets which are paired are of similar size, then WPGMA is essentially UPGMA.[9]

As an example consider $N = \{v_1, \ldots, v_5\}$ and the (ultrametric) distance matrix given by

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ |
|---|---|---|---|---|---|
| $v_1$ | 0 | 8 | 4 | 6 | 8 |
| $v_2$ | | 0 | 8 | 8 | 4 |
| $v_3$ | | | 0 | 6 | 8 |
| $v_4$ | | | | 0 | 8 |
| $v_5$ | | | | | 0 |

UPGMA creates the classification

$$\mathcal{C} = \{\{v_1\}, \{v_2\}, \{v_3\}, \{v_4\}, \{v_5\}, \{v_1, v_3\}, \{v_1, v_3, v_4\}, \{v_2, v_5\}, N\}. \tag{10.15}$$

The related unrooted $N$-tree is

$$T = (((v_1 v_3) v_4)(v_2 v_5)), \tag{10.16}$$

which is the correct tree and of length 15.

[9] Strangely, sometimes in referenced works, our method UPGMA is called WPGMA and vice versa.

# 11

# THE PHYLOGENY

> Nothing in biology makes sense except in the light of evolution.
>
> Theodosius Dobzhansky

As it became accepted that evolution was to be understood in terms of Mendelian genetics and Darwinian natural selection, so too it became clear that this understanding could not be sought only at a qualitative level. A fundamental problem is the reconstruction of species' evolutionary past, which is called the phylogeny of those species. Here, trees are widely used to represent these relationships.

## 11.1 PHYLOGENETIC TREES

> The holy grail of phylogenetics is the reconstruction of the one true tree of life.
>
> J.T.Thorley and R.D.M.Page

The underlying principle of phylogeny is to try to group "living entities" according to their level of similarity.
In biology for example, such trees ("phylogenies") typically represent the evolutionary history of a collection of extant species or the line of descent of some

gene. No two members of a species are exactly the same - each has slight modifications from their parents. As environmental conditions change, nature will favour that branch of a species with some particular modification; as time goes on another mutation of the basic stock will become dominant. In this way, all species are continually evolving. This evolution occurs in a number of ways at the same time: some species die out and some become new species in their own right. This was already seen by Darwin [25]:

> The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth. The green and budding twigs may represent existing species; and those produced during each former year may represent the long succession of extinct species... The limbs divided into great branches, and these into lesser and lesser branches, were themselves once, when the tree was small, budding twigs; and this connexion of the former and present buds by ramifying branches may well represent the classification of all extinct and living species in groups subordinate to groups... From the first growth of the tree, many a limb and branch has decayed and dropped off, and these lost branches of various sizes may represent those whole orders, families, and genera which have now no living representatives, and which are known to us only from having been found in a fossil state... As buds give rise by growth to fresh buds, and these, if vigorous, branch out and overtop on all a feebler branch, so by generation I belive it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever branching and beautiful ramifications.

Historically, this was a new idea: The concept of species having a continuity through time was only developed in the late 17th century; higher life forms were no longer thought to transmute into different kinds during the lifetime of an individual. It took over 150 years from the development of this concept before a rooted tree was proposed by Darwin.[1]

The phylogenetic tree can therefore be thought of as a central metaphor for evolution, providing a natural and meaningful way to order data, and with an enormous amount of evolutionary information contained within its branches.

[1] Note that in Darwin's fundamental book *The origin of species* [25] there is exactly one figure, and this shows the description of the evolutionary history by a tree.

Clearly, this idea is attractive, but how are we to find the tree? Note that there are several difficulties, even in the definition of the problem:

- What is the tree of life? A tree which is given by a classification or the evolutionary tree?
- What is the mechanism of evolution? Darwin provided mutation and natural selection, which suggested a scientific model for the relation of species.
- Darwin's evolutionary tree is neither obvious, nor easy to find.
- There must be some criterion for deciding which of the many phylogenies that may be drawn most closely resembles the actual evolutionary changes.
- Darwin saw another difficulty in the underlying problems. In a letter to Huxley he wrote: "The time will come, I believe, though I shall not live to see it, when we shall have fairly true genealogical trees of each great kingdom of Nature."  
  Hence, it seems impossible to describe the "Great Darwin Tree" since the diversity of the living world is staggering: more than two million existing species of plants and animals have been named and described; many more - both existing and past - remain to be discovered.
- Considering the origin of life: Was there just one, or more than one "starting point"? What does we know about the last universal common ancestor, if it exists?
- It has been argued that the "Tree of Life" is perhaps really a "Web of Life", as mechanisms such as hybridization, recombination and swapping of genes probably play a role in evolution.

A nice representation of this subject has been given in [26], [64], and [86]. A survey about *What Evolution is* was given by Mayr [59]. For the history of Darwin's theory compare [10] and [89].

Each species can be described in terms of a sequence of specific values, called characters. These characters were originally morphological, that is derived from an analysis of an organism's form and structure, but how are these values measurable?  
In biology, "characters" describe attributes of the species under consideration and are the data that biologists typically use to reconstruct phylogenetic trees. We wish to consider characters for species in a morphological sense. To do this we assume that there is given a (finite or infinite) state space $\mathcal{C}$ of characters.

We also assume that there is a metric in $\mathcal{C}$. Discrete character data are those for which a function $f$ assigns a character state $f_{ij}$ to each taxon $i$ for each character $j$.

As sequence data became readily available it was predicted an end to this conflict. Now, the biological units are written in words constructed from the letters corresponding either to amino acids, which generate proteins, or to nucleotides forming DNA or RNA molecules. By comparing such words one can construct evolutionary (phylogenetic) trees showing how closeness of the words in the tree corresponds to the closeness of the unit. In other words,

**The Phylogenetic Tree Problem**
**Given:** A set of sequences, each representing a taxon.
**Find:** Their phylogenetic tree, representing its evolutionary history.

The set of leaves represent the given taxa, the internal vertices are the ancestors, and the root of the tree represents the common ancestor of all. The phylogenetic tree of life shows when groups of organisms arose and gives the basic relationships between them.

First, molecular sequence data was used by Fitch and Margoliash in their landmark paper [33] dealing with cytochrome c sequences. The basic idea in that field is that species (given by their sequences) which appear to be closely related should have diverged more recently than species which appear to be less closely related.
This task is more complicated than it seems at first glance; Gould [38] wrote:

> When systematists, also known as taxonomists, set out to reconstruct the phylogeny (evolutionary history) of a group of species that they think are related, they have before them the species living today and the fossil record. To reconstruct a phylogenetic history as closely as possible, they must make inferences based on observational and experimental data. The difficulty is that what can be measured is *similarity*, whereas the goal is to determine *relatedness*.

Note that the definition of similarity cannot be the problem of the mathematical analysis. This is, in any case, the task of the biological sciences. But mathematics can help to check if the choice was not false.
Overviews of tree making algorithms are given in [22], [42], [51], [65], and [83]

## 11.2 THE PERFECT PHYLOGENY PROBLEM

We introduce a character-based approach to reconstructing evolutionary history. The input is a set of attributes called characters that objects may possess. The most important problem in morphological phylogenetics is selecting the characters. Here opposing side picking out is the favourite method. On the other hand, characters must be coded if there are more than two distinct possibilities. The basic assumptions regarding characters are:

- The characters being considered are "meaningful" in the context of phylogenetic tree reconstruction.
- The characters can be inherited independently from one another.
- All observed states for a given character should have evolved from one "original state" of a common ancestor of the objects.[2]

Note that character in this context does not refer to a member of an alphabet; for simplicity we will use natural numbers for character states.

A taxon $v$ over a set $C$ of $m$ characters is a vector $v \in \mathbb{N}^m$. $c(v)$ is the state of $v$ on character $c$ or the state of $c$ for $v$. $A_c$ is the set of allowed states for $c(v)$, assuming that $A_c = \{0, \ldots, r_c - 1\}$ for some integer $r_c \geq 2$.
Let $N = \{v_1, \ldots, v_n\}$ be a set of $n$ taxa, represented by an $n \times m$ character-state matrix $M = (f_{ij})$, where $f_{ij}$ is the state of taxon $v_i$ on character $j$.
A $N$-tree $T = (V, E)$ should represent the phylogeny for $N$, with internal vertices (which may also be labelled) representing hypothetical ancestors to the given taxa, where

(i) Each of the taxa labels exactly one leaf of $T$, and vice versa;

(ii) Each of the characters labels exactly one edge of $T$, but not necessarily vice versa; and

(iii) For any taxon $v$, the characters that label the edges along the unique path from the root to $v$ describe the character states of $V$.

A character $c$ is called convex in a $N$-tree if for every $f \in A_c$, the set of vertices $\{v \in V : c(v) = f\}$ induces a subtree of $T$. An $N$-tree $T$ is called a perfect

[2] Characters that obey this assumption are called homologous.

phylogeny if every $c \in C$ is convex in $T$.
The interpretation of such a tree for $M$ is that it gives an estimate of the evolutionary history of the taxa, based on the following biological assumptions:

(i) The root of the tree represents an ancestral taxon that has none of the present $m$ characters.

(ii) Each of the characters change from one state to another state exactly once and never changes back to the zero state.[3]

**The Perfect Phylogeny Problem**
**Given:** A set of taxa on a set of characters, represented by a character-state matrix.
**Determine:** Whether a perfect phylogeny exists.

Steel [79] showed that the perfect phylogeny problem is very hard in the sense of computational complexity.

We will now restrict ourselve to the binary case, that is we allow a character to take exactly two states: $M$ is a $0-1$-matrix. Here, we will see that the problem can solved efficiently.
For any column $k$ of $M$, let $O_k$ be the set of taxa with a 1 in column $k$- that is the taxa that have character $k$. The major fact and the basis for an efficient solution of the (binary) perfect phylogeny problem is

**Theorem 11.2.1** *The matrix $M$ has a phylogenetic tree if and only if for every pair of columns $i$ and $j$, either $O_i$ and $O_j$ are disjoint or one contains the other.*

This theorem is intuitively clear, and a complete proof is given in [41] and [76]. To make this technique clearer, Gusfield [41] furnishes the following small example: Let $M_1$ be the matrix

$$\begin{array}{c|ccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\hline
v_1 & 1 & 1 & 0 & 0 & 0 \\
v_2 & 0 & 0 & 1 & 0 & 0 \\
v_3 & 1 & 1 & 0 & 0 & 1 \\
v_4 & 0 & 0 & 1 & 1 & 0 \\
v_5 & 0 & 1 & 0 & 0 & 0
\end{array}$$

[3] Hence any taxon below that edge definitely have that character.

and let $M_2$ be

$$\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline v_1 & 1 & 1 & 0 & 0 & 0 \\ v_2 & 0 & 0 & 1 & 0 & 1 \\ v_3 & 1 & 1 & 0 & 0 & 1 \\ v_4 & 0 & 0 & 1 & 1 & 0 \\ v_5 & 0 & 1 & 0 & 0 & 1 \end{array}$$

$M_1$ has a phylogenetic tree, namely $(((v_1v_3)v_5)(v_2v_4))$, but matrix $M_2$ not.

## 11.3 DISTANCE METHODS

The phylogenetic analysis of a family of (related) nucleic acid or protein sequences is the determination of how the family might have been derived during evolution. Evolutionary relationships among the sequences are depicted by placing the sequences on the leaves of a tree. The branching relationships on the internal vertices of the tree then reflect which sequences are related. Starting with a set of known present-day objects a phylogenetic tree may be constructed by first assigning each object a leaf of the tree and then assigning ancestral and unknown objects to the internal nodes. Roughly speaking, we have the following relationship:

| Level | | |
|---:|:---:|:---:|
| In taxonomy | OTU = operational taxonomic unit | HTU = hypothetical taxonomic unit |
| Species/genes | extant | extinct |
| Placement in time | existing unit | ancestor |
| Classification | individuals | class |
| Vertex in the tree | leaf | internal vertex |

The general idea is the following: Let $N = \{v_1, \ldots, v_n\}$ be a set of individuals (OTU's). We assume that $N$ is embedded in a metric space $(X, \rho)$, such that we represent the distances between the members of $N$ by a symmetric distance matrix

$$D = (d_{ij}) = (\rho(v_i, v_j)), \tag{11.1}$$

where $i, j = 1, \ldots, n = |N|$.
We would like to build a phylogenetic tree for $N$. If we fix a $N$-tree $T = (V, E)$ we obtain a tree-metric $\rho_T$.[4] The broad aim of distance methods is to determine a (or all) tree(s) $T$ for which $\rho_T$ is as close as possible to $\rho$.

Consider the following example. For $n = 3$ we have only one $N$-tree $T$ with one internal vertex $w$. Searching the edge-lengths $l_i = \rho_T(v_1, w)$ such that $\rho = \rho_T$ is solving the following system

$$\begin{aligned} l_1 + l_2 &= d_{12} \\ l_1 + l_3 &= d_{13} \\ l_2 + l_3 &= d_{23}, \end{aligned}$$

which is given by

$$\begin{aligned} l_1 &= \frac{1}{2}(d_{12} + d_{13} - d_{23}) \\ l_2 &= \frac{1}{2}(d_{12} + d_{23} - d_{13}) \\ l_3 &= \frac{1}{2}(d_{13} + d_{23} - d_{12}). \end{aligned}$$

(Note that the values on the right-hand side are non-negative due to the triangle inequality.) Hence, we have a unique tree which reflects the phylogeny with respect to given distances.

A collection of distances $D = (d_{ij}) = (\rho(v_i, v_j))$ for $i, j = 1, \ldots, n = |N|$ is called additive if there is a $N$-tree $T$ such that

$$d_{ij} = \rho_T(v_i, v_j). \tag{11.2}$$

We saw that for $n$ each distances are additive. In general, this is not true for $n \geq 4$. We characterize additve distances by the following theorem.

**Theorem 11.3.1** *(The four-point condition)*
*$(d_{ij})_{i,j=1,\ldots,n}$ is additive if and only if for every set of four distinct numbers $1 \leq i, j, k, l \leq n$ two of the three[5] sums $d_{ij} + d_{kl}$, $d_{ik} + d_{jl}$ and $d_{il} + d_{jk}$ coincide and are greater than or equal to the third one.*

[4] Remember that this is the length of a shortest path.

[5] compare 3.4.

The proof is constructive and given by the so-called neighbor-joining algorithm which is well known from our pair group methods, but with a slight modification: the choice of the pair of taxa is to unite is not directly derived from the distances; it is modified the a value which estimates the length of the internal edges. More exactly:

**Algorithm 11.3.2** *Let $N = \{v_1, \ldots, v_n\}$ be a set of individuals (OTU's) with given distances $(d_{ij})_{i,j=1,\ldots,n}$.*

1. *For $i = 1$ until $n$ do*
$$r_i = \frac{1}{n-2} \sum_{k=1}^{n} d_{ik};$$
2. *Pick up a pair $i, j$ for which*
$$d_{ij} - (r_i + r_j)$$
*is minimal;*
3. *Group together $v_i$ and $v_j$ to form $v_{n+1}$;*
*(the OTU $v_{n+1}$ represents an internal vertex of the future tree)*
*Compute a new distance from $v_{n+1}$ to each of the other $v_k$ by*
$$d_{n+1k} = \frac{d_{ik} + d_{jk} - d_{ij}}{2};$$
4. *Repeat the procedure until there is only one OTU.*

## 11.4 THE MAXIMUM PARSIMONY METHOD

The maximum parsimony method is a popular technique for reconstructing phylogenetic trees from sequences or states of characters.
The principle of Maximum Parsimony involves the identification of a combinatorial structure that requires the smallest number of evolutionary changes. This is an application of Ockham's razor, according to which the best hypoth-

esis is the one requiring the smallest number of assumptions.[6,7] It means that among all possible structures we seek one which satisfies only one, namely the condition of length minimization.
While the validity of parsimony has been debated, it can be justified on biological grounds, see [63]. But note several problems with this point of view:

- The amount of change recovered by parsimony is, by definition, the smallest possible amount that is consistent with the data. The actual amount of evolutionary change may have been somewhat larger.
- We may find more than one minimal-length tree interconnecting the given entities.

Parsimony is widely used in practice, as attested by the popularity of software such as PAUP, which stands for "Phylogenetic analysis using parsimony"; see [42] and [82].

**Algorithm 11.4.1** *(Fitch [34]) Let $N$ be a set of $n$ sequences in a sequence space $(A^d, \rho_H)$: $N = \{v_k = v_{k,1}, \ldots, v_{k,d} : k = 1, \ldots, n\}$, and let a binary $N$-tree $T = (V, E)$ be given. Then do:*

```
1. For each position $i = 1, \ldots, d$ do
   1. Label each leaf $v_k$ with $\{v_{k,i}\}$;
```

[6] Or in other words:

(a) It is futile to do with more what can be done with fewer.

(b) More precisely in Latin: Entia non sunt multiplicanda praeter necessitatem.

(c) More roughly spoken: Keep it simple.

This is true, but not in a simple sense. Cavalli-Sforza [20]:

> ... it does not necessarily follow that a method of tree reconstruction minimizing the number of mutations is the best or uses all the information contained in the sequences. The minimization of the number of mutations is intuitively attractive because we know that mutations are rare. There may be some confusion, however, between the advantage of minimizing the number of mutations and sometimes invoked parallel of Ockham's razor ..., which was developed in the context of medieval theology. The extrapolation of Ockham's razor to the number of mutations in an evolutionary tree is hardly convincing.

Note that in this case minimizing the number of assumptions does not mean minimizing the number of mutations, or the steps of an evolution, it means that among all possible network structures we seek one which satisfies only few conditions. With the "razor", Ockham cuts out all superfluous, redundant explanations.

[7] For a broader philosophical discussion of Ockham's razor see [12] and [70], [71].

```
    $L_i := 0$;
    2. Until all vertices are marked do
    Find an unlabelled vertex which is adjacent to two marked vertices with the marks $N_1$ and $N_2$;
    Mark the unmarked vertices with
    (a) $N_1 \cap N_2$ if $N_1 \cap N_2 \neq \emptyset$; otherwise
    (b) $N_1 \cup N_2$ and $L_i := L_i + 1$;
2. $L(T) := \sum_{i=1}^{d} L_i$.
```

The correctness of Fitch's algorithm is proven in [44]. In particular, it is shown that the final answer is independent of the vertices chosen when moving through the tree.
The algorithm easily computes the length of the tree. On the other hand, there are exponentially many binary trees. Hence, the Fitch algorithm 11.4.1 uses a great amount of time.

After applying 11.4.1 we have marks for all the internal vertices in the tree. However, some marks have more than one letter and hence are ambiguous. There are several methods for choosing which one of the possible states yields the most parsimonious reconstruction; the simplest one is Farris' method: go back up the tree assigning to any internal vertex that is ambiguous the intersection of its mark with that of its immediate ancestor.

## 11.5 CONSENSUS TREES

A consensus tree summarises information common to two or more trees. In other words:

- A phylogenetic tree summarizes phylogenetic information;
- A consensus tree summarizes the information in a set of trees.
  Here, we have two additional observations:
  (a) We can combine heterogeneous data, and
  (b) We can find hidden phylogenetic information.

For instance, Cavalli-Sforza [16] compares the species tree and the tree of languages for human populations. This gives many hints for the prehistoric development of mankind.

Consensus trees are helpful to taxonomists: When they had completed a classification, it may that with data from another source the classification is different, or that by using a different clustering method a new classification results. The taxonomist may wish to form an overall classification which takes account of the information shared in each classification, however it is obtained.

**The Consensus Tree Problem**
**Given:** A collection $T_i$ of $N$-trees, $i = 1, \dots, m$.
**Find:** An $N$-tree summarizing the phylogenetic information of all $T_i$.

There are many ways to combine $N$-trees into a single tree; see [75].[8] The methods differ in what aspect of tree information they use, and how frequently that information must be shared among the trees to be included in the consensus. The most commonly used are the strict consensus and the majority-rule consensus trees.

Suppose that $T_1, \dots, T_m$ are $N$-trees. Each of the trees has the same leaves, namely the members of $N$. We are interested in an $N$-tree $T$ described by one of the following methods.

The strict consensus tree includes only those splits that occur in all the trees. That means

$$\mathcal{S}(T) = \bigcap_{i=1}^{m} \mathcal{S}(T_i). \tag{11.3}$$

We can relax the requirement that a split of $T$ occur in all trees, and instead retain those splits occuring in a majority of the trees.

**Algorithm 11.5.1** *For each of the $N$-trees $T_1, \dots, T_m$, mark the vertices inductively as follows:*

1. *Mark the leaf $v$ with $\{v\}$;*
2. *If the vertices $v_1, \dots, v_r$ have been marked with $N_1, \dots, N_r$ and $v$ is the common ancestor of $v_1, \dots, v_r$, then mark $v$ with $N_1 \cup \dots \cup N_r$.*

*The consensus tree $T$ consists of exactly those vertices whose mark occurs in more than half of the $T_i$.*

[8] For simple but fundamental limitations on consensus tree methods compare [80].

# REFERENCES

[1] M. Aigner. *Diskrete Mathematik.* Vieweg, 1993.

[2] S.F. Altschul. A Protein Alignment Scoring System Sensitive at All Evolutionary Distances. *J. Molecular Evolution*, 36:290–300, 1993.

[3] I. Anderson. *A First Course in Discrete Mathematics.* Springer, 2001.

[4] T.K. Attwood and D.J. Parry-Smith. *Introduction to bioinformatics.* Prentice Hall, 1999.

[5] F.J. Ayala. The myth of Eve - molecular-biology and human origins. *Science*, 270:1930–1936, 1995.

[6] F.J. Ayala and A.A. Escalaute. The evolution of human populations: A molecular perspective. *Mol. Phyl. Evol.*, 5:188–201, 1996.

[7] H.-J. Bandelt, P. Forster, B.C. Sykes, and M.B. Richards. Mitochondrial Portraits of Human Populations Using Median Networks. *Genetics*, 141:743–753, 1995.

[8] R. Bellman. *Dynamic Programming.* Princeton University Press, 1957.

[9] C. Berge. *Graphs.* Elsevier Science Publishers, 1985.

[10] P.J. Bowler. *Evolution: The history of an idea.* Univ. Calif. Press, 1984.

[11] N.F. Britton. *Essential Mathematical Biology.* Springer, 2003.

[12] J.R. Brown. *philosophy of mathematics.* Routledge, 1999.

[13] R.L. Cann and A.C. Wilson. Models of human evolution. *Science*, 217:303–304, 1982.

[14] R.M. Cann, M. Stoneking, and A. Wilson. Mitochondrial DNA and Human Evolution. *Nature*, 325:31–36, 1987.

[15] J.L. Casti. *Five More Golden Rules.* John Wiley & Sons, 2000.

[16] L.L. Cavalli-Sforza. Stammbäume von Völkern und Sprachen. In B. Streit, editor, *Evolution des Menschen*, pages 118–125. Spektrum Akademischer Verlag, 1995.

[17] L.L. Cavalli-Sforza. *Gene, Völker und Sprachen.* Carl Hanser Verlag, 1999.

[18] L.L. Cavalli-Sforza and F. Cavalli-Sforza. *Verschieden und doch gleich.* Knaur, 1996.

[19] L.L. Cavalli-Sforza and A.W.F. Edwards. Phylogenetic analysis: models and estimation procedures. *Evolution*, 21:550–570, 1967.

[20] L.L. Cavalli-Sforza, P. Menozzi, and A. Piazza. *The History and Geography of Human Genes.* Princeton University Press, 1994.

[21] A. Cayley. A theorem on trees. *Quart. Math.*, 23:376–378, 1889.

[22] P. Clote and R. Backofen. *Computational Molecular Biology.* John Wiley & Sons, 2000.

[23] B. Comrie, S. Matthews, and M. Polinsky. *The Atlas of Languages.* Quarto Publishing Plc., 1996.

[24] J.H. Conway and R.K. Guy. *The Book of Numbers.* Springer, 1996.

[25] C. Darwin. *The Origin of Species.* London, 1859.

[26] P. Davies. *The Fifth Miracle.* Penguin, 1998.

[27] M.O. Dayhoff. Atlas of Protein Sequence and Structure. Technical Report 5, National Biomedical Research Foundation, Washington, D.C., 1978.

[28] E.W. Dijkstra. A note on two problems in connection with graphs. *Numer. Math.*, 1:269–271, 1959.

[29] A. Dress and R. Wetzel. The Human Organism - a Place to Thrive for the Immuno-Deficiency Virus. In E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, and B. Burtschy, editors, *New Approaches in Classification and Data Analysis*, pages 636–643. Springer Verlag, 1994.

[30] M. Eigen. *Stufen zum Leben.* Serie Piper, 1992.

[31] B.S. Everitt. *Cluster Analysis.* Arnold, 1993.

[32] M. Farkas. *Dynamical Models in Biology.* Academic Press, 2001.

[33] W. Fitch and E. Margoliash. Construction of Phylogenetic Trees. *Science*, 155:279–284, 1967.

[34] W.M. Fitch. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology*, 20:406–416, 1971.

[35] A. Gierer. *Die gedachte Natur: Ursprünge der modernen Wissenschaft.* rowohlt, 1998.

[36] E.N. Gilbert. Gray codes and paths on the n-cube. *Bell System Tech. J.*, 37:815–826, 1958.

[37] M. Glaubrecht. *Die ganze Welt ist eine Insel.* Hirzel Verlag, 2002.

[38] J.L. Gould and W.T. Keeton. *Biological Sciences.* W.W.Norton and Company, 1996.

[39] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison-Wesley, Boston, 1989.

[40] D. Graur and W.H. Li. *Fundamentals of Molecular Evolution.* Sinauer Associates, Inc., 1999.

[41] D. Gusfield. *Algorithms on Strings, Treees, and Sequences.* Cambridge University Press, 1997.

[42] B.G. Hall. *Phylogenetic Trees Made Easy.* Sinauer Associates, Sunderland, MA, 2001.

[43] F. Harary and E.M. Palmer. *Graphical Enumeration.* Academic Press, 1973.

[44] J.A. Hartigan. Minimum mutation fits to a given tree. *Biometrics*, 29:53–65, 1973.

[45] M. Hendy, C.H.C. Little, and D. Penny. Comparing trees with pendant vertices labelled. *SIAM J. Appl. Math.*, 44:1054–1065, 1984.

[46] A. Isaev. *Introduction to Mathematical Methods in Bioinformatics.* Springer, 2004.

[47] D. Johanson, L. Johanson, and B. Edgar. *Ancestors: In Search of Human Origins.* Villard Books, 1994.

[48] D. Jungnickel. *Graphen, Netzwerke und Algorithmen.* BI Wissenschaftsverlag, Mannheim, 1994.

[49] M. Kanehisa. *Post-genome Informatics.* Oxford University Press, 2000.

[50] R.M. Karp. Reducibility among combinatorial problems. In R.E. Miller and J.W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103, New York, 1972.

[51] V. Knoop and K. Müller. *Gene und Stammbäume.* Spektrum, 2006.

[52] T.W. Körner. *The Pleasures of Counting.* Cambridge University Press, 1996.

[53] T.W. Körner. *Mathematisches Denken.* Birkhäuser, 1998.

[54] B. Korte and J. Vygen. *Combinatorial Optimization.* Springer, 2000.

[55] J.B. Kruskal. On the shortest spanning subtree of a graph and the travelling salesman problem. *Proc. of the Am. Math. Soc.*, 7:48–50, 1956.

[56] V.I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Phys. Dokl.*, 10:707–710, 1966.

[57] B. Lomborg. *The sceptical environmentalist.* Cambridge University Press, 2002.

[58] J. Matoušek and J. Nešetřil. *Diskrete Mathematik.* Springer, 2002.

[59] E. Mayr. *What Evolution is.* Basic Books, New York, 2001.

[60] P Morrison and P. Morrison. *Powers of Ten.* Scientific American Books, 1982.

[61] S. Olson. *Mapping Human History - Discovering the Past Through Our Genes.* Houghton Mifflin Company, 2002.

[62] R. Otter. The Number of Trees. *An. Math.*, 49:583–599, 1948.

[63] R.D.M. Page and E.C. Holmes. *Molecular Evolution: A Phylogenetic Approach.* Blackwell Science, 1998.

[64] E. Pennisi. Modernizing the Tree of Life. *Science*, 300:1692–1697, 2003.

[65] D. Penny and M. Hendy. Phylogenetics: Parsimony and Distance Methods. In D.J.Balding et al., editor, *Handbook of Statistical Genetics*, pages –. John Wiley & Sons, Ltd., 2000.

[66] H. Prüfer. Ein neuer Beweis eines Satzes über Permutationen. *Arch. Math. Phys.*, 27:742–744, 1918.

[67] F.S. Roberts. *Graph Theory and its Applications to Problems of Society.* SIAM, 1978.

[68] D. Robinson and L.R. Foulds. Comparison of phylogenetic trees. *Math. Biosci.*, 53:131–147, 1981.

[69] S.M. Ross. *Probability Models for Computer Science.* Harcourt Academic Press, 2002.

[70] B. Russell. *A History of Western Philosophy.* George Allen & Unwin, 1945.

[71] B. Russell. *Philosophie des Abendlandes.* Europa Verlag, 1950.

[72] V.M. Sarich and A.C. Wilson. Immunological time scale for hominoid evolution. *Science*, 158:1200–1203, 1967.

[73] K.-H. Schleifer and M. Horn. Mikrobielle Vielfalt - die unsichtbare Biodiversität. *Biologie heute*, 6:1–5, 2000.

[74] R.-H. Schulz. *Codierungstheorie.* Vieweg, 1991.

[75] C. Semple and M. Steel. *Phylogenetics.* Oxford University Press, 2003.

[76] J. Setubal and J. Meidanis. *Introduction to Computational Molecular Biology.* PWS Publishing Company, 1997.

[77] T.F. Smith, M.S. Waterman, and W.M. Fitch. Comparative Biosequence Metrics. *J. Molecular Evolution*, 18:38–46, 1981.

[78] H.J. Sprengel and O. Wilhelm. *Funktionen und Funktionalgleichungen.* Deutscher Verlag der Wissenschaften, 1984.

[79] M. Steel. The complexity of reconstructing trees from qualitative characters and subtrees. *Journal of Classification*, 9:91–116, 1992.

[80] M. Steel, A.W.M. Dress, and S. Böcker. Simple but Fundamental Limitations on Supertree and Consensus Tree Methods. *Syst. Biology*, 49:363–368, 2000.

[81] I. Stewart. *Galois Theory.* Chapman and Hall, 1998.

[82] D.L. Swofford. *PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods (software).* Sinauer Associates, Sunderland, MA, 2000.

[83] D.L. Swofford and G.J. Olsen. Phylogeny Reconstruction. In D.M. Hills and C. Moritz, editors, *Molecular Systematics*, pages 411–501. Sinauer Associates, 1990.

[84] V.V. Vazirani. *Approximation Algorithms.* Springer, 2001.

[85] M. Vingron, H.-P. Lenhof, and P. Mutzel. Computational Molecular Biology. In M. Dell'Amico, F. Maffioli, and S. Martello, editors, *Annotated Bibliographies in Combinatorial Optimization*, pages 445–471. John Wiley and Sons, 1997.

[86] P.D. Ward and D. Brownlee. *Rare Earth.* Springer, 2000.

[87] M.S. Waterman. Sequence Alignments. In M.S. Waterman, editor, *Mathematical Methods for DNA-Sequencing*, pages 53–92. CRC Press, 1989.

[88] M.S. Waterman. Applications of Combinatorics to Molecular Biology. In R.L. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, pages 1983–2001. Elsevier Science B.V., 1995.

[89] T.P. Weber. *Darwin und die Anstifter.* Du-Mont, 2000.

[90] A.C. Wilson and R.L. Cann. Afrikanischer Ursprung des modernen Menschen. In B. Streit, editor, *Evolution des Menschen*, pages 86–93. Spektrum Akademischer Verlag, 1995.

[91] H. Yockey. *Information Theory and Molecular Biology.* Cambridge University Press, 1992.