IT University of Copenhagen

Foundations of Computing – Discrete Mathematics MSc

Re-exam

6 March, 2018

Instructions (Read Carefully)

Contents: The exam contains 14 questions for a total of 100 points. The exam is divided into two parts: The first part has 10 multiple choice questions and the second part has 4 open ended questions.

What to check: In the multiple-choice questions, there is one and only one correct answer. You should only check 1 box.

Definitions and theorems: At the end of this document (page 12) you can find some definitions and theorems that could be useful for answering some of the questions.

Info about you: Write *clearly* your full name and your date of birth (DoB) on *every page* (top-right).

—**IMPORTANT**—

Only information written on the pages 1–11 will be evaluated.
Anything else that you hand in will NOT be considered for the final evaluation!

Part I. Answer the following multiple choice questions.

S **1.** (6 pts) Which of the following statements is **true**?

  $\boxed{\text{A}}$ $1 \subseteq \{1, 2, 3\}$

  $\boxed{\checkmark}$ $\{1\} \times \{2, 3\} = \{(1, 2), (1, 3)\}$

  $\boxed{\text{C}}$ $\{1\} \in \{1, 2, 3\}$

  $\boxed{\text{D}}$ $\mathcal{P}(\{1, 2, 3\}) = \{\emptyset, 1, 2, 3\}$

F **2.** (6 pts) Let the two functions $f : \mathbb{N} \to \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$ be defined as follows:

$$f(n) = n + 1 \qquad\qquad \text{for all } n \in \mathbb{N}$$
$$g(n) = 2 \cdot n \qquad\qquad \text{for all } n \in \mathbb{N}$$

  where $\mathbb{N}$ is the set $\{0, 1, 2, \dots\}$. Which of the following statements is **true**?

  $\boxed{\text{A}}$ f is onto.

  $\boxed{\text{B}}$ g is a one-to-one correspondence.

  $\boxed{\checkmark}$ f is one-to-one.

  $\boxed{\text{D}}$ $(g \circ f)(n) = 2 \cdot n + 1$ for all $n \in \mathbb{N}$.

  Solution: If $f(n) = f(m)$, then $n + 1 = m + 1$. By subtracting 1 from both sides of the latter equation we obtain that $n = m$. Hence, f is one-to-one.

N **3.** (6 pts) Which of the following is **true** for all positive integers a, b and q?

  $\boxed{\text{A}}$ If $ab \equiv 0 \pmod{q}$ then either $a \equiv 0 \pmod{q}$ or $b \equiv 0 \pmod{q}$.

  $\boxed{\text{B}}$ If $a^2 \equiv a \pmod{q}$ then $a \equiv 1 \pmod{q}$.

  $\boxed{\checkmark}$ If $a \equiv a + b \pmod{q}$ then $b \equiv 0 \pmod{q}$.

  $\boxed{\text{D}}$ If $a^2 \equiv 0 \pmod{q}$ then $a \equiv 0 \pmod{q}$.

  Solution: $a \equiv a + b \pmod{q}$ implies by Theorem 3 that $q \mid (a - (a + b))$ and thus $q \mid -b$. Consequently, $q \mid b$, which means that $b \equiv 0 \pmod{q}$.

CP **4.** (6 pts) A toy factory has two machines A and B. Machine A produces 75% of the factory output and machine B produces 25% of the factory output. 2% of the toys produced by machine A are defective and 6% of those produced by B are defective. If a toy is chosen at random from the factory output and turns out to be defective, what is the likelihood it was produced by machine B.

$\boxed{\text{A}}$ 70%

$\boxed{\text{B}}$ 50%

$\boxed{\text{C}}$ 40%

$\boxed{\text{D}}$ 22.33%

Solution: Let $P(B)$ be the probability that a toy is produced by machine B and $P(A)$ be the probability that a toy is produced by machine A. Let $P(D)$ be the probability that a toy is defective. We want to calculate the conditional probability $P(B|D)$. Bayes' formula say that

$$P(B|D) = \frac{P(D|B)P(B)}{P(D)}$$

But we have

$$P(D) = P(D|A)P(A) + P(D|B)P(B) = 0.02 \cdot 0.75 + 0.06 \cdot 0.25 = 0.03$$

Now

$$P(B|D) = \frac{0.06 \cdot 0.25}{0.03} = \frac{0.015}{0.03} = 0.5$$

CP **5.** (6 pts) Consider 5 people A, B, C, D and E. In how many ways can we arrange them in a queue such that B and C are not in a row (i.e. B and C are separated by at least one other person)?

A 24

B 48

✔ 72

D 96

Solution: There are 5! possible ways to arrange A,B,C,D, and E in a queue. There are 4! possible ways to arrange the same five people in a queue such that B is always behind C. Similarly, There are 4! possible ways to arrange the same five people in a queue such that C is always behind B. By the difference rule there are $5! - 2 \cdot 4!$ ways to arrange those five people in a queue such that B and C are not in a row. That is, $5! - 2 \cdot 4! = 5 \cdot 4! - 2 \cdot 4! = 3 \cdot 4! = 3 \cdot 24 = 72$

L **6.** (6 pts) One of the compound propositions below is logically equivalent to the compound proposition $\sim q \wedge p$. Which one?

A $(q \wedge p) \wedge \sim q$

B $q \vee \sim(\sim p \vee q)$

✔ $(q \vee p) \wedge \sim q$

D $\sim(\sim q \vee \sim p)$.

Solution:

$$(q \vee p) \wedge \sim q$$
$$\equiv (q \wedge \sim q) \vee (p \wedge \sim q)$$
$$\equiv \mathbf{f} \vee (p \wedge \sim q)$$
$$\equiv p \wedge \sim q$$
$$\equiv \sim q \wedge p$$

R

7. (6 pts) Let $R = \{(0,0), (0,1), (1,3), (2,1)(2,2), (2,3), (3,0)\}$ be a relation on the set $S = \{0,1,2,3\}$. Which of the following statements is **true**?

$\boxed{A}$ R is reflexive and transitive, but not antisymmetric.

$\boxed{B}$ R is antisymmetric, but neither reflexive nor transitive.

$\boxed{C}$ R is transitive and antisymmetric, but not reflexive.

$\boxed{D}$ R is reflexive, but neither transitive nor antisymmetric.

A   **8.** (6 pts) Which one of the regular expressions given below defines a language that **does not** contain the string *abbab*?

$\boxed{A}$ $(a|b)^*$

$\boxed{B}$ $(a|b)(b|ba)^*(ab|\varepsilon)$

$\boxed{C}$ $(a|b)^*(a|bb)$

$\boxed{D}$ $b^*(ab|ba)^*b^*$

A **9.** (6 pts) Consider the following grammar:

$$S \rightarrow AB$$
$$S \rightarrow \varepsilon$$
$$A \rightarrow Sa$$
$$B \rightarrow bS$$

Which one of the strings below belongs to the language generated by this grammar?

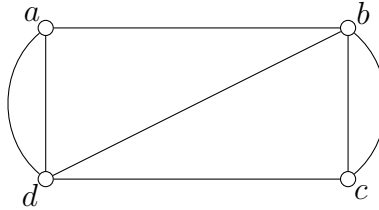$\boxed{A}$ *baba*

$\boxed{\checkmark}$ *abab*

$\boxed{C}$ *ababa*

$\boxed{D}$ *aabb*

Solution: *abab* can be generated as follows :

$$S \Rightarrow AB \Rightarrow SaB \Rightarrow SabS \Rightarrow Sab \Rightarrow ABab \Rightarrow SaBab \Rightarrow SabSab \Rightarrow Sabab \Rightarrow abab$$

G   **10.** (6 pts) Let G be the following graph:

Which of the following statements is **true**?

$\boxed{A}$ G is not connected.

$\boxed{B}$ G has an Euler circuit.

✔ G has an Euler trail.

$\boxed{D}$ G is a simple graph.

Solution: Both a and c have an odd degree (namely 3) all other vertices have an even degree. Hence, there is an Euler trail from a to c.

Part II. Answer the following questions. Be brief but precise. Your correct use of mathematical notation is an important aspect of your answer.

11. (12 pts) Show that $(q \wedge p) \wedge (p \to \sim(p \vee q))$ is a contradiction by constructing a truth table for all subexpressions.

Solution:

p	q	$q \wedge p$	$p \vee q$	$\sim(p \vee q)$	$p \to \sim(p \vee q)$	$(q \wedge p) \wedge (p \to \sim(p \vee q))$
T	T	T	T	F	F	F
T	F	F	T	F	F	F
F	T	F	T	F	T	F
F	F	F	F	T	T	F

p	q	$q \wedge p$	$p \vee q$	$\sim(p \vee q)$	$p \to \sim(p \vee q)$	$(q \wedge p) \wedge (p \to \sim(p \vee q))$
T	T	T	T	F	F	F
F	T	F	T	F	T	F
T	F	F	T	F	F	F
F	F	F	F	T	T	F

L **12.** (12 pts) Let the sequence $a_0, a_1, a_2, \ldots$ be given by the following recursive definition

$$a_k = 2a_{k-1} + 1 \qquad \text{for all } k \geq 1$$
$$a_0 = 1$$

Prove by mathematical induction that $a_n = 2^{n+1} - 1$ for all $n \geq 0$.

Solution: We want to prove the statement

$$a_n = 2^{n+1} - 1 \qquad\qquad (P(n))$$

for all $n \geq 0$.

Basis step: Let $n = 0$. We have

$$2^{0+1} - 1 = 2^1 - 1 = 2 - 1 = 1$$

and, by definition $a_0 = 1$. Hence, the basis step is verified.

Inductive step:
Suppose that $k \geq 0$ and that $P(k)$ holds, that is,

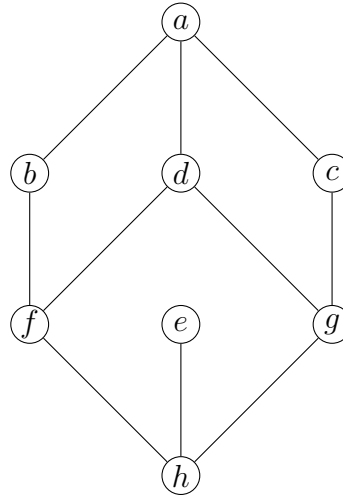$$a_k = 2^{k+1} - 1 \qquad\qquad \text{(inductive hypothesis)}$$

We must show that $P(k+1)$ holds, that is,

$$a_{k+1} = 2^{(k+1)+1} - 1 \qquad\qquad (P(k+1))$$

The following calculation shows that $P(k+1)$ holds:

$$
\begin{aligned}
a_{k+1} &= 2a_k + 1 & \text{(by definition of the sequence)} \\
&= 2(2^{k+1} - 1) + 1 & \text{(inductive hypothesis)} \\
&= 2 \cdot 2^{k+1} - 2 + 1 \\
&= 2 \cdot 2^{k+1} - 1 \\
&= 2^{(k+1)+1} - 1
\end{aligned}
$$

S **13.** (8 pts) Let $(\{a, b, c, d, e, f, g, h\}, \preceq)$ be the partially ordered set defined by the following Hasse diagram:

(a) Find all minimal elements.

Solution: h

(b) Find all maximal elements.

Solution: a, e

(c) Find all least elements.

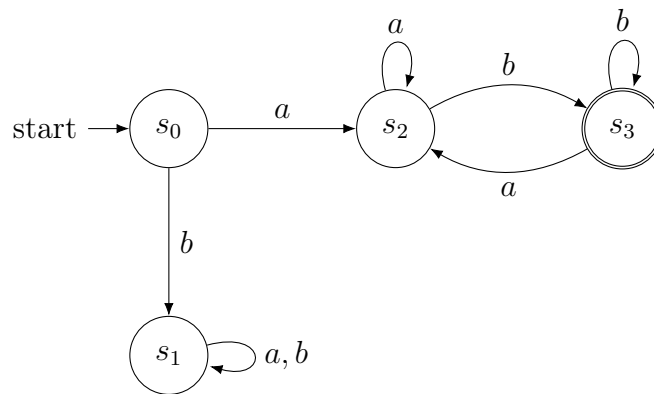*Solution*: h is the least element.

(d) Find all greatest elements.

Solution: There is no greatest element.

A **14.** (8 pts) Construct a finite-state automaton A with input alphabet $\{a, b\}$ that accepts the set of all strings that start with a and end with b. That is, A must satisfy

$$L(A) = \{awb \mid w \in \{a, b\}^*\}$$

Describe the automaton A using a next-state table *or* a transition diagram.

Solution:

Definitions and theorems

Logic

The truth table for a number of logical operators is given below.

p	q	$\sim p$	$p \vee q$	$p \wedge q$	$p \rightarrow q$	$p \leftrightarrow q$
T	**T**	**F**	**T**	**T**	**T**	**T**
T	**F**	**F**	**T**	**F**	**F**	**F**
F	**T**	**T**	**T**	**F**	**T**	**F**
F	**F**	**T**	**F**	**F**	**T**	**T**

A compound proposition is called a *tautology* if it is always true no matter what the truth values of the propositional variables are. A compound proposition that is always false is called a *contradiction*.

The compound propositions p and q are called *logically equivalent* if $p \leftrightarrow q$ is a tautology. The notation $p \equiv q$ denotes that p and q are logically equivalent.

Given any statement variables p, q, and r, a tautology **t** and a contradiction **c**, the following logical equivalences hold.

1.	*Commutative laws:*	$p \wedge q \equiv q \wedge p$	$p \vee q \equiv q \vee p$
2.	*Associative laws:*	$(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$	$(p \vee q) \vee r \equiv p \vee (q \vee r)$
3.	*Distributive laws:*	$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$
4.	*Identity laws:*	$p \wedge \mathbf{t} \equiv p$	$p \vee \mathbf{c} \equiv p$
5.	*Negation laws:*	$p \vee \sim p \equiv \mathbf{t}$	$p \wedge \sim p \equiv \mathbf{c}$
6.	*Double negative law:*	$\sim(\sim p) \equiv p$	
7.	*Idempotent laws:*	$p \wedge p \equiv p$	$p \vee p \equiv p$
8.	*Universal bound laws:*	$p \vee \mathbf{t} \equiv \mathbf{t}$	$p \wedge \mathbf{c} \equiv \mathbf{c}$
9.	*De Morgan's laws:*	$\sim(p \wedge q) \equiv \sim p \vee \sim q$	$\sim(p \vee q) \equiv \sim p \wedge \sim q$
10.	*Absorption laws:*	$p \vee (p \wedge q) \equiv p$	$p \wedge (p \vee q) \equiv p$
11.	*Negations of* **t** *and* **c***:*	$\sim\mathbf{t} \equiv \mathbf{c}$	$\sim\mathbf{c} \equiv \mathbf{t}$

Sets

A *set* is an (unordered) collection of objects, called *elements* or *members*. We write $a \in A$ to indicate that a is an *element of* A

A set A is a *subset* of a set B, written $A \subseteq B$, if for all a, $a \in A$ implies $a \in B$.

Two sets A and B are equal if $A \subseteq B$ and $B \subseteq A$.

The *power set* of a set A, denoted $\mathcal{P}(A)$, is the set of all subsets of A. That is, $\mathcal{P}(A) = \{B \mid B \subseteq A\}$.

The *union* of two sets A and B is the set $A \cup B = \{x \mid x \in A \vee x \in B\}$.

The *intersection* of A and B is the set $A \cap B = \{x \mid x \in A \wedge x \in B\}$.

The *difference* of two sets A and B is the set $A - B = \{x \mid x \in A \wedge x \notin B\}$.

The *complement* of a set A is the set $A^c = \{x \in \mathcal{U} \mid x \notin A\}$, where $\mathcal{U}$ is the *universal set*.

The *Cartesian product* of A and B is the set $A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$. That is, $A \times B$ is the set of all pairs (a, b) with $a \in A$ and $b \in B$.

Functions

Given two sets A and B, a *function* f from A to B (written $f : A \to B$) is an assignment of exactly one element of B to each element of A.

A function $f : A \to B$ is *onto* if for every element $b \in B$ there is an element $a \in A$ such that $f(a) = b$.

A function $f : A \to B$ is *one-to-one* if $f(a) = f(b)$ implies $a = b$ for all a and b in the domain of f.

A function f is a *one-to-one correspondence* if it is both one-to-one and onto.

Given two functions $g : A \to B$ and $f : B \to C$, the *composition* of f and g (written $f \circ g : A \to C$) is given by

$$(f \circ g)(a) = f(g(a)) \qquad \text{for all } a \in A$$

Let $f : X \to Y$. If $f(x) = y$ we say that y is the *image* of x and x is an *inverse image* (or *preimage*) of y. The *range* (or *image*) of f is the set of all values of f, i.e.

$$\mathsf{range}(f) = \{y \in Y \mid \exists x \in X . f(x) = y\}$$

The *inverse image* of an element $y \in Y$ is the set of all inverse images of y, i.e.

$$\text{inverse image of } y = \{x \in X \mid f(x) = y\}$$

Relations

Let A and B be sets. A *binary relation* from A to B is a subset of $A \times B$.

A *relation on a set* A is a binary relation from A to A.

A relation R on a set A is called *reflexive* if $(a, a) \in R$ for every element $a \in A$.

A relation R on a set A is called *symmetric* if $(b, a) \in R$ whenever $(a, b) \in R$, for all $a, b \in A$.

A relation R on a set A is called *antisymmetric* if whenever $(a, b) \in R$ and $(b, a) \in R$, then $a = b$, for all $a, b \in A$,.

A relation R on a set A is called *transitive* if whenever $(a, b) \in R$ and $(b, c) \in R$, then $(a, c) \in R$, for all $a, b, c \in A$.

The *transitive closure* of a relation R is the smallest transitive relation T such that $R \subseteq T$.

A relation is an *equivalence relation* if it is reflexive, symmetric and transitive.

A relation is a *partial order relation* if it is reflexive, antisymmetric and transitive.

A *partially ordered set* is a pair (A, R) consisting of a set A and a partial order relation R on A.

Elements a and b of a partially ordered set $(A, \preceq)$ are said to be *comparable* if $a \preceq b$ or $b \preceq a$. Otherwise, a and b are said to be *incomparable*.

Let $(A, \preceq)$ be a partially ordered set. An element $a \in A$ is called

- a *greatest element* of A, if $x \preceq a$ for all $x \in A$

- a *least element* of A, if $a \preceq x$ for all $x \in A$

- a *maximal element* of A if, for all $x \in A$, either $x \preceq a$ or a and x are incomparable

- a *minimal element* of A if, for all $x \in A$, either $a \preceq x$ or a and x are incomparable

Number Theory

Given two integers a and b, with $a \neq 0$, we say that a *divides* b (written $a \mid b$) if there exist an integer c such that $b = ac$, or equivalently, if $\frac{b}{a}$ is an integer.

Theorem 1. *Let a, b, and c be any integers.*

1. *If $a \mid b$ and $a \mid c$ then $a \mid (b + c)$.*

2. *If $a \mid b$ then $a \mid bc$.*

3. *If $a \mid b$ and $b \mid c$ then $a \mid c$.*

Theorem 2 (The Quotient/Remainder Theorem). *Given any integer a and a positive integer d, there exist unique integers q and r such that*

$$a = dq + r \qquad and \ 0 \leq r < d$$

In Theorem 2 the value d is called the *divisor*, a is the *dividend*, q is the *quotient*, and r is the *remainder*. Then div and mod are defined as a div $d = q$, a mod $d = r$. Remember that the remainder cannot be negative.

Let n and m be any integers and d a positive integer. We write $n \equiv m \pmod{d}$ if

$$(n \bmod d) = (m \bmod d)$$

Theorem 3. *Let n and m be any integers and d a positive integer then*

$$n \equiv m \pmod{d} \quad \textit{if and only if} \quad d \mid (n - m)$$

Let n be a nonnegative integer and $b_r b_{r-1}...b_0$ a finite sequence of binary digits, i.e. $b_i \in \{0, 1\}$ for $0 \leq i \leq r$. The sequence $b_r b_{r-1}...b_0$ is a binary representation of n if

$$n = b_r \cdot 2^r + b_{r-1} 2^{r-1} + ... + b_1 \cdot 2 + b_0$$

For example 101101 is the binary representation of $1 \cdot 2^5 + 0 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2 + 1 = 45$

The *greatest common divisor* of two integers a and b, not both zero, is denoted by $\gcd(a, b)$ and is the largest integer that both divides a and divides b.

The *least common multiple* of two positive integers a and b is denoted by $\mathrm{lcm}(a, b)$ and is the smallest positive integer c such that $a \mid c$ and $b \mid c$.

Theorem 4. *If a and b integer, not both zero, then*

$$\gcd(a, b) = \gcd(a, a \bmod b)$$

Counting

The table below states how many different ways there are to order n distinct objects, and how many ways there are to choose r objects out of n distinct objects (depending on whether the order matters or repetition is allowed).

Order n objects (permutation)	$P(n) = n! = n \cdot (n-1) \ldots 2 \cdot 1$	
Choose r objects from n	without repetition	with repetition
- order matters (r-permutation)	$P(n, r) = \frac{n!}{(n-r)!}$	n^r
- order doesn't matter (r-combination)	$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$	$\binom{n+r-1}{r}$

Theorem 5 (Binomial Theorem). *Given real numbers a and b, and non-negative integer n,*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Let A be a finite set. We write $N(A)$ to denote the number of elements in A. We say A has $N(A)$ elements or A is a set of size $N(A)$

Theorem 6. *If A and B are finite sets, then*

$$N(A \cup B) = N(A) + N(B) - N(A \cap B)$$

Theorem 7 (Difference Rule). *If A is a finite set and B is a subset of A, then*

$$N(A - B) = N(A) - N(B).$$

Probability

A *sample space* is the set of all possible outcomes of a random process or experiment. An *event* is a subset of a sample space. The probability of an event E is denoted $P(E)$. If S is a finite sample space in which all outcomes are equally likely and E is an event in S, then the probability of E is

$$P(E) = \frac{N(E)}{N(S)}$$

Let S be a sample space. A *probability function P* is a function from the set of all events in S to the set of real numbers satisfying

1. $0 \leq P(A) \leq 1$ for all events A in S.

2. $P(\emptyset) = 0$ and $P(S) = 1$.

3. If A and B are disjoint events in S, i.e. $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

Theorem 8. *If A is any event in a sample space S then the probability of the complement event $A^c = S - A$ is*

$$P(A^c) = 1 - P(A)$$

Theorem 9. *If A and B are events in a sample space S then*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Suppose the sample space of an experiment or a random process is given by $S = \{a_1, ..., a_n\}$ where a_i is a real number for all $1 \le i \le n$. Suppose that each a_i occur with probability p_i for $1 \le i \le n$. The *expected value* of the process is

$$\sum_{i=1}^{i=n} a_i p_i = a_1 p_1 + a_2 p_2 + ... + a_n p_n$$

Let A and B be events in a sample space S. If $P(A) \ne 0$, then the *conditional probability of B given A*, denoted $P(B|A)$, is

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Theorem 10 (Bayes' Theorem). *Suppose that a sample space S is the union of mutually disjoint events $B_1, ..., B_n$ where $P(B_i) \ne 0$ for all $1 \le i \le n$. Suppose that A is an event in S and that $P(A) \ne 0$. Then*

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ... + P(A|B_n)P(B_n)}$$

Graphs & Trees

A *graph* is a triple (V, E, f) consisting of: a finite, nonempty set V of vertices; a finite set E of edges; and a function $f : E \to \{\{x, y\} \mid x, y \in V\}$.

An edge with only one endpoint is called a *loop*. Two edges with the same set of endpoints are called *parallel*. Vertices are called *adjacent* if they are endpoints of the same edge (if the edge is a loop, its endpoint is called *adjacent to itself*). We say that an edge is *incident on* its endpoints. Two edges with a common endpoint are called *adjacent*. A vertex on which no edge is incident is called *isolated*.

A graph is called *simple* if it has no loops or parallel edges.

Let G be a graph and v a vertex of G. The *degree of v*, denoted $\deg(v)$, is the number of edges that are incident on v, with an edge that is a loop counted twice. The *total degree of G* is the sum of the degrees of all vertices in G.

Theorem 11 (Handshake Theorem). *A graph with m edges has total degree $2m$.*

The table below gives an overview over the different kinds of walks in a graph:

	repeated edge	repeated vertex	same start & end vertex	must contain at least 1 edge
Walk	allowed	allowed	allowed	no
Trail	no	allowed	allowed	no
Path	no	no	no	no
Closed walk	allowed	allowed	yes	no
Circuit	no	allowed	yes	yes
Simple Circuit	no	first and last only	yes	yes

A *subgraph* of G is a graph (V', E', f') with $V' \subseteq V$, $E' \subseteq E$ and $f'(e) = f(e)$ for all $e \in E'$.

A graph (V, E, f) is *connected* if there exists a walk from u to v for all $u, v \in V$. A *connected component* of a graph G is a maximal connected subgraph C of G (i.e. every connected subgraph of G is either a subgraph of C or has no common vertices with C).

An *Euler circuit* of a graph G is a circuit that contains every edge and every vertex of G.

An *Euler trail* of a graph G is a trail that contains every edge and every vertex of G.

Theorem 12. *A graph has an* Euler circuit *if and only if it is connected and every vertex has positive even degree.*

Theorem 13. *Given two distinct vertices u and v in a graph G, there exists an Euler trail from u to v if and only if G is connected, u and v have odd degree, and all other vertices have positive even degree.*

A graph is called *circuit-free* if it has no circuits. A *tree* is a connected and circuit-free graph. In a tree, a vertex is called a *leaf* if it has degree 1, and an *internal vertex* if it has degree 2 or greater. A tree which consists of only one vertex is called *trivial*.

Theorem 14. *Let n be a positive integer. A tree with n vertices has $n - 1$ edges.*

Automata, Regular Expressions, Grammars

An *alphabet* Σ is a finite set of *symbols*.

A *string* (or *word*) w over Σ is a finite sequence of symbols from Σ.

A *language* A is a set of strings over some Σ.

ε is the *empty string*.

The *length* of w, written $|w|$, is the number of symbols that w contains. We write $|w|_a$ for the number of times the symbol a occurs in w.

Given words $w_1 = a_1 a_2 \ldots a_m$ and $w_2 = b_1 b_2 \ldots b_n$, we write $w_1 w_2$ to mean their *concatenation*: $a_1 a_2 \ldots a_m b_1 b_2 \ldots b_n$.

The concatenation of two languages L_1 and L_2 is defined as the language

$$L_1 L_2 = \{w_1 w_2 \mid w_1 \in L_1 \text{ and } w_2 \in L_2\}$$

The *Kleene closure* of a language L is the language $L^* = \{\varepsilon\} \cup L \cup LL \cup LLL \cup \ldots$

A *finite-state automaton* is a 5-tuple (S, I, N, s_0, F) consisting of: a finite set of states S; a finite set I (the input alphabet); a transition function $N : S \times I \to S$; an initial state s_0; and a set $F \subseteq S$ of accepting states.

The language accepted by an automaton A is denoted by $L(A)$:

$$L(A) = \{w \mid w \in I^* \text{ and } N^*(s_0, w) \in F\}$$

where $N^* : S \times I^* \to S$ is defined by

$$N^*(s, \varepsilon) = s$$
$$N^*(s, aw) = N^*(s', w) \qquad \text{where } s' = N(s, a)$$

A *regular expression* r over an alphabet Σ can be built using: $\emptyset$; ε; a symbol $a \in \Sigma$; concatenation $r\,s$ and union $r \mid s$ of regular expressions r, s; and Kleene closure r^* of a regular expression r.

The language $L(r)$ defined by a regular expression r is

$$L(\emptyset) = \emptyset \qquad\qquad L(a) = \{a\} \qquad\qquad L(r \mid t) = L(r) \cup L(r)$$
$$L(\varepsilon) = \{\varepsilon\} \qquad\qquad L(r\,t) = L(r)\,L(r) \qquad\qquad L(r^*) = L(r)^*$$

A *grammar* $G = (V, T, S, P)$ consists of: A set of symbols V called *vocabulary*; a set $T \subset V$ of terminal symbols (symbols in $N = V - T$ are called *non-terminal*); a starting symbol $S \in V$; and a set P of productions of the form $z_0 \to z_1$, where $z_0, z_1 \in V^*$ and z_0 must contain at least one non-terminal symbol.

Given a grammar G and two strings $s, t \in V^*$, we write $s \Rightarrow t$, if s can be rewritten to t by applying one production from P. We write $s \overset{*}{\Rightarrow} t$, and say that t is derivable from s, if s can be rewritten to t by applying several productions from P. The language defined by the grammar G is the set of strings over T derivable from S:

$$L(G) = \left\{w \mid w \in T^* \text{ and } S \overset{*}{\Rightarrow} w\right\}$$