Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

Moja záverečná práca

(šablóna)

Bakalárska práca

2018   **Bc. Janko Hraško PhD.**

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

Moja záverečná práca
(šablóna)

Bakalárska práca

Študijný program:	Informatika
Študijný odbor:	9.2.1. Informatika
Školiace pracovisko:	Katedra počítačov a informatiky (KPI)
Školiteľ:	Leslie Lamport
Konzultant:	Donald E. Knuth

Košice 2018   **Bc. Janko Hraško PhD.**

Abstrakt v SJ

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kľúčové slová v SJ

LaTeX, programovanie, sadzba textu

Abstrakt v AJ

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Kľúčové slová v AJ

LaTeX, programming, typesetting

Bibliografická citácia

HRAŠKO, Janko. *Moja záverečná práca (šablóna)*. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2018. **??**s. Vedúci práce: Leslie Lamport

Tu vložte zadávací list pomocou príkazu
\thesisspec{cesta/k/suboru/so/zadavacim.listom}
v preambule dokumentu.
Kópiu zadávacieho listu skenujte čiernobielo (v odtieňoch sivej) na 200 až 300
DPI! Nezabudnite do jednej práce vložiť originál zadávacieho listu!

Čestné vyhlásenie

Vyhlasujem, že som záverečnú prácu vypracoval(a) samostatne s použitím uvedenej odbornej literatúry.

Košice, 13.5.2018 ....

Vlastnoručný podpis

Poďakovanie

Na tomto mieste by som rád poďakoval svojmu vedúcemu práce za jeho čas a odborné vedenie počas riešenia mojej záverečnej práce.

Rovnako by som sa rád poďakoval svojim rodičom a priateľom za ich podporu a povzbudzovanie počas celého môjho štúdia.

V neposlednom rade by som sa rád poďakoval pánom *Donaldovi E. Knuthovi* a *Leslie Lamportovi* za typografický systém LaTeX, s ktorým som strávil množstvo nezabudnuteľných večerov.

Obsah

Úvod

The order of DNA sequence and its variations are the very aspect which dictates the developmental processes of an organism, determines susceptibility to various diseases and uniquely identifies each creature. This area has always been on the periphery of the interests of scientific society, since the discovery in 1869 by Swiss-born biochemist Fredrich Miescher. For instance, The Human Genome Project (HGP) which started on October 1, 1990 and completed in April 2003 was one of the greatest feats of exploration in history of science. It was aimed at reading all the DNA sequences of our species, Homo sapiens. All in all, HGP introduced us the ability to read nature's complete genetic blueprint for building a human being. However, despite the successful completion of the project, a number of unknown DNA properties is still exists and demands the profound studying.

The knowledge of the genome structure has significantly increased in the past few decades thanks to the recent developments in the field of advanced analyzing techniques. The Sanger sequencing technology has been traditionally used to elucidate the DNA sequencing information since it was developed in the 1977th. However, it is capable of obtaining sequences of maximum length of 800 base pairs per one operation, which makes the sequencing process much longer and complicated. In spite of development of new sequencing techniques, some technology limits exist. For instance, the human genome in particular presents a number of major obstacles to correct read alignment, due to its size (3 GB) and complexity (48% repetitive sequences), as do other plant plant and vertebrate genomes.

In addition, it is impossible to assemble the whole genome sequence of species using the data merely of one individual due to occurrence of the single nucleotide polymorphisms and mutations which affect the precise result. Several sequencing algorithms and searching methods were developed to deal with such issues which are the basis of the bioinformatics. To be precise, the science was developed to deal with the next problems: assembling the complete nucleic acid sequence from the smaller parts, its comparison, analyzing and searching of similarities.

The usual eukaryotic genome consists not only of nuclear DNA, but also of

DNA which is isolated from it and belongs to some organelles (mitochondrial mDNA, plastid DNA) that became a part of the cells in the evolution process. To identify key features and determine the exact genes at the complete DNA sequence, to distinguish the segments belonging to particular chromosomes it must visualized in some way. The whole genome might be visualized either as the two dimensional representation of the nucleotide sequence or as the 3D model of the spatial DNA or RNA architecture. The first way allows to analyze each gene and precisely identify each protein that it encodes and to trace the kinship of species, while the second way provides us with the possibility of understanding the inner cellular processes and the very interaction between different enzymes and nucleic acid from the chemical point of view. This work concerns mainly the first method of visualizing sequencing data.

Although several DNA processing tools exist, the problem of representing different genome properties which might vary at various species, concerning either the number of particular genes or complete chromosomes (if they are present), remains still actual. Moreover, the processing of the genome and its visualization demand an efficient approach, concerning the size of data and computational capabilities of an average computer. This work aims at representing some key genome properties in such a way.

Formulácia úlohy

Text záverečnej práce musí obsahovať sekciu s formuláciou úlohy resp. úloh riešených v rámci záverečnej práce. V tejto časti autor rozvedie spôsob, akým budú riešené úlohy a tézy formulované v zadaní práce. Taktiež uvedie prehľad podmienok riešenia.

1 Analytická časť

In most eukaryotic and prokaryotic organisms the hereditary material is either linear double-stranded DNA (deoxyribonucleic acid) molecules or a circular double-stranded DNA molecule. However, some extracellular life forms, might use RNA (ribonucleic acid) as the building block for their genome. For instance, viruses have a genome composed of either single-stranded DNA, double-stranded DNA or RNA, depending on the type of a virus. Therefore, a genome itself, is the complete content of genetic information in an organism, or in other words, all the unique DNA or RNA sequences the organism possesses.

1.1 Nucleotides

Both of DNA and RNA are polymeric molecules, that are composed of linear chains of various combinations of four different subunits, called nucleotides. The nucleotide itself is the basic unit of the DNA and RNA molecules, the monomer, which, however, could be found in the cell not only as the bearer of the genetic information, but also as a carrier of energy used to power enzymatic reactions. A five-carbon-atom sugar, a phosphate group and a nitrogenous base are three distinct components which, combined together, make up the quite complex nucleotide molecule. The combination of sugar and base is called a nucleoside, while the phosphate-sugar-base is termed a nucleotide. The nucleotide bases can be either a single-ringed pyrimidine or a double-ringed purine. Dinucleotide, trinucleotide and polynucleotide are the terms corresponding to two, three or many nucleotides connected with each other respectively.

A nucleotide can be either a purine or pyrimidine. Guanine (G) and adenine (A) are the common purines for both of DNA and RNA; the pyrimidine called cytosine (C) is also present in both nucleic acids. However, the pyrimidine uracil (U) is limited only to RNA, being replaced with thymine (T) in DNA. There are merely two base-pair combinations that are permissible – A base-paired with T (U) and C base-paired with G. It happens due to the geometries of the nucleotide
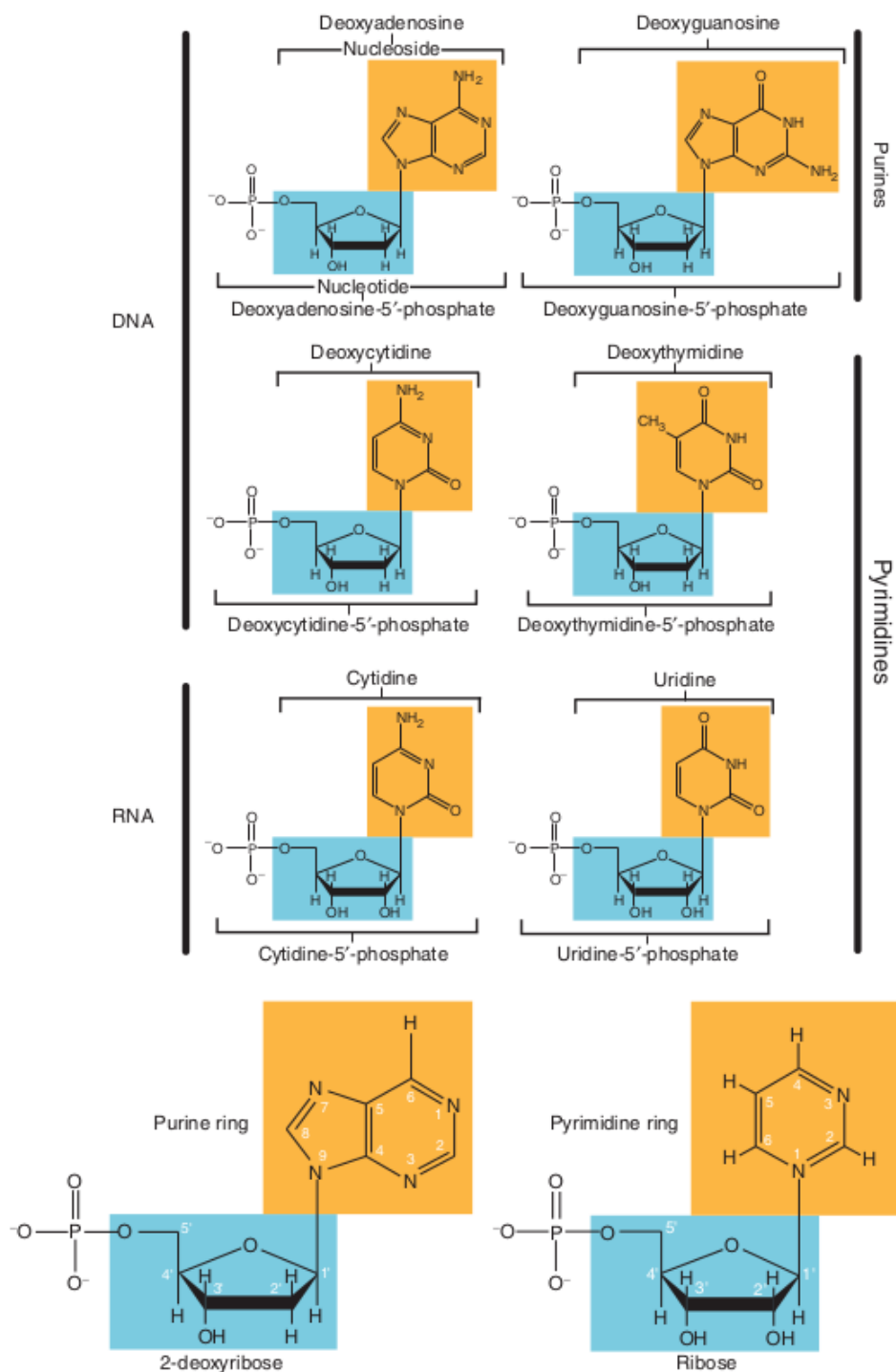
Obr. 1.1: The structures of the pyrimidines and purines found in DNA and RNA. The sugar groups are highlighted in blue and the nitrogenous bases are highlighted in orange. The atoms of the sugar are numbered from 1 to 5. The atoms of the purine ring are numbered from 1 to 9, while those of the pyrimidine ring are numbered from 1 to 6.

bases and relative positions of atoms which participate in the connection. This property makes two sequences of polynucleotides in helix complement. Discrete

nucleotides are attached to each other through sugar–phosphate bonds that connect the phosphate group on the 5′ carbon of one nucleotide with the hydroxyl group on the 3′ carbon of another nucleotide. The base pairing between adenine and thymine (uracil) involves two hydrogen bonds, while between cytosine and guanine involves three hydrogen bonds.

1.2   Nucleodic acid spatial stucture

As the three-dimensional structure of a nucleotide is not completely rigid, it is possible for DNA to have various spatial architectures: A-form, B-form, Z-form and the circular one. The position of the base relatively to the five-carbon-atom sugar can be changed by a rotation around the N-glycosidic bond and, in this way, significantly affect the three dimensional configuration of the molecule and helix consequently.

Tabuľka 1.1: DNA double helix

Features of the different conformations of the DNA double helix			
Feature	B-DNA	A-DNA	Z-DNA
Type of helix	Right-handed	Right-handed	Left-handed
Number of base pairs per turn	10	11	12
Distance between base pairs (nm)	0.34	0.29	0.37
Distance per complete turn (nm)	3.4	3.2	4.5
Diameter (nm)	2.37	2.55	1.84
Major groove	Wide, deep	Narrow, deep	Flat
Minor groove	Narrow, shallow	Wide shallow	Narrow, deep

Moreover, although usually single-stranded, some RNA sequences have the ability to form a double helix. However, double helix RNA is rare and has nothing in common with the genome itself, since only the single-stranded RNA molecules appear to participate in some genome related processes in the eukaryotic and prokaryotic organisms. Since circular DNA may exist in several forms inclu-

ding single-stranded c-DNA, intact double-stranded c-DNA (closed circles with both strands covalently linked), nicked ds-c-DNA (only one strand covalently linked) and "concatenated circles" their properties are not described in the following table.

1.3 Chromosomes in eukaryotic genomes

In eukaryotic cells nucleic acid is situated in a membrane-bound organelle called the nucleus. The nuclear genome is split into a set of linear double-helix DNA molecules, each contained in a chromosome. No exceptions to this pattern are known: all eukaryotes that have been studied have at least two chromosomes and the DNA molecules are always linear[?]. The only variability at this level of organization of eukaryotic genome is coherent with the number of chromosomes. Moreover, it appears, that biological features of an organism have no dependence on the number of chromosomes.
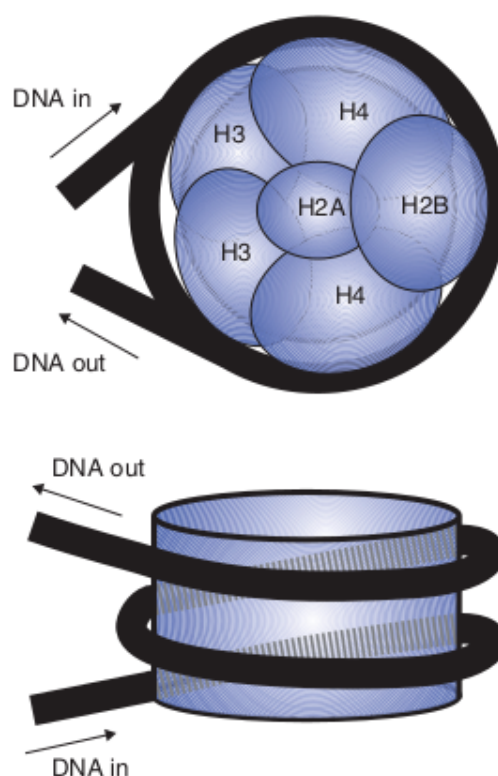
Obr. 1.2: The nucleosome structure. H2A, H2B, H3 and H4 represent different types of histones.

Despite the size of a nucleus (5-10 um)[?], an overall length of DNA in the human cell is approximately 2.1m[?] and can be packed inside the cell because
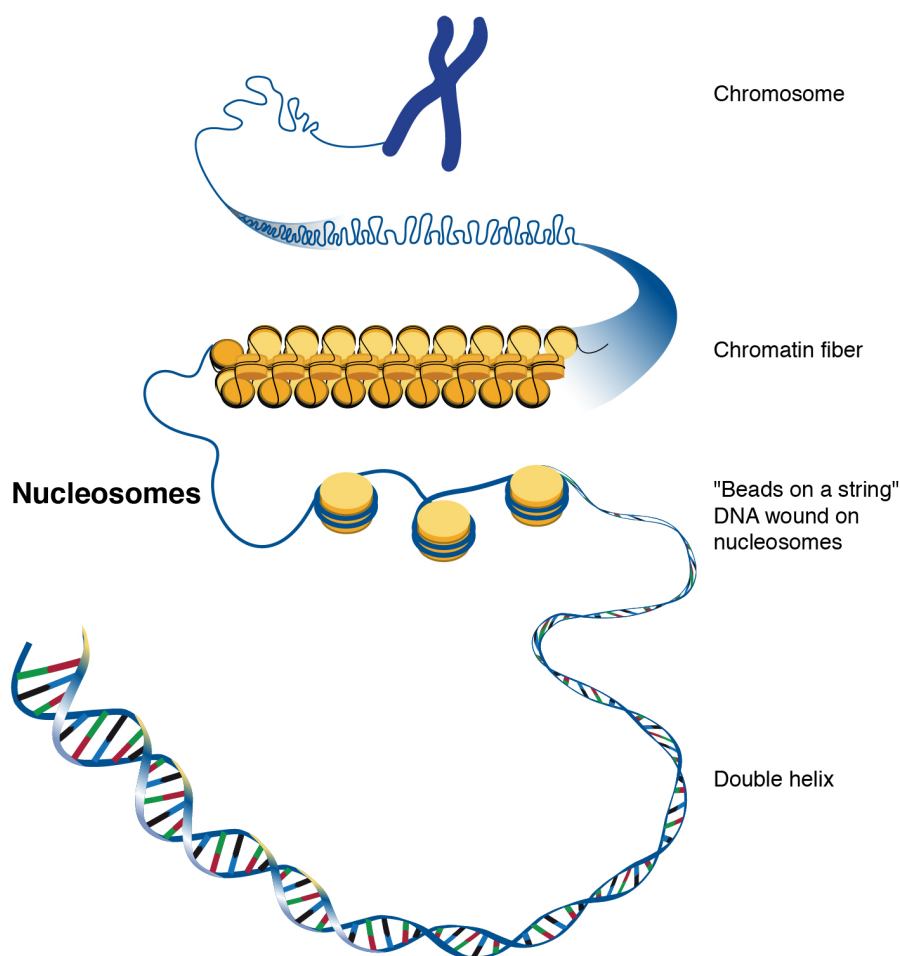
of the method the nucleic acid is stored. The genetic material in viruses and bacteria consists of strings of DNA or RNA almost devoid of proteins. However, in eukaryotes, a substantial quantity of protein is associated with the DNA to form chromatin. At the lowest level, the DNA is organized by wrapping DNA strands around he proteins called histones, that contain a large amount of positively charged amino acids arginine and lysine. Those amino acids, and histones in general, play the crucial structural role, making it possible to bind the negative charged phosphate groups of the DNA nucleotides.

Averagely, the DNA rolled around the histones consists of 140-150 base pair, dependently on the species[?]. Such a complex of DNA and histones is termed a nucleosome. These nucleosomes can be further coiled into increasingly larger coils up until forming chromosomes [8]. However, tight coiling of DNA limits cells ability to access DNA and to process it. Instead of being constantly coiled, the nucleic acid is usually found in a state called chromatin where some segments of acid are tightly reeled (heterochromatin), while other segments are entirely open (euchromatin). Euchromatin DNA is is highly accessible by the molecular complexes used by the cell and therefore is easier to manipulate with.

The amount and extent of packing are determined by a sell, to control which sections of the genome can be expressed and which cannot. It affects cellular function and appears to be the predominant cause of differentiating cells type, while having the same DNA.

1.4   Lorem ipsum

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Obr. 1.3: Ncleosomes as the part of a chromosome.

1.5 Aliquam eu malesuada urna

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

- v knihe [**book**] autor prezentuje naozaj odvážne myšlienky

- nemenej zaujímavé výsledky publikuje ďalší autor v článku [**article**]

- v konferenčnom príspevku [**conference**] sú uvedené tiež zaujímavé veci

- LaTeX[1] je typografický jazyk

Given a set of numbers, there are elementary methods to compute its Greatest Common Divisor, which is abbreviated GCD. This process is similar to that used for the Least Common Multiple (LCM).

1.5.1   Donec vehicula consequat

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.
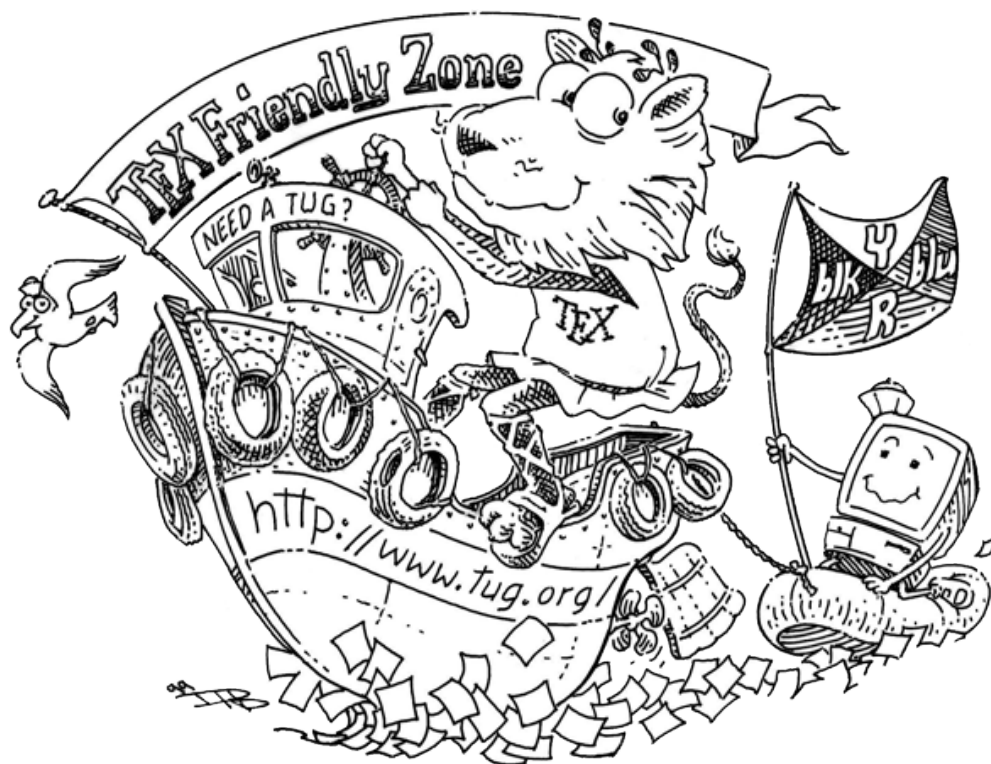
1.5.2   Nullam in mauris consectetur

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Zdrojový kód 1.1: Program, ktorý pozdraví celý svet

```c
#include <stdio.h>
int main() {
```

---

[1] https://www.latex-project.org/

Obr. 1.4: LATEX Friendly Zone

```
/* Print Hello, World! */
printf("Hello, World!\n");
return 0;
}
```

1.5.3   Vestibulum tristique elementum varius

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Tabuľka 1.2: Country list

Country List			
Country Name or Area Name	ISO ALPHA 2 Code	ISO ALPHA 3 Code	ISO numeric Code
Afghanistan	AF	AFG	004
Aland Islands	AX	ALA	248
Albania	AL	ALB	008
Algeria	DZ	DZA	012
American Samoa	AS	ASM	016
Andorra	AD	AND	020
Angola	AO	AGO	024

1.6 Phasellus id pretium neque

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisissem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi necante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultriciesvel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero utmetus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit ametante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

2 Syntetická časť

Syntetická časť opisuje metódy použité na syntézu riešenia a opisuje syntézu samotného riešenia (zvyčajne je to návrh/implementácia softvérového resp. hardvérového riešenia), pričom sa opiera o závery analytickej časti práce. Začína od toho, ako sa bude riešenie používať: najdôležitejšie scenáre používania a používateľské rozhranie, ktoré bude tieto scenáre efektívne podporovať. Až potom je na rade vnútorná architektúra alebo použité technológie. Syntetická časť tvorí zvyčajne ½ jadra práce.

Syntetickú časť práce vhodne rozdeľte do kapitol a pomenujte ich podľa toho, čomu sú venované.

3 Vyhodnotenie

Vyhodnocovacia časť je kľúčovou časťou záverečnej práce. Tato časť obsahuje vyhodnotenie navrhnutého (vytvoreného) riešenia. Uprednostňované je objektívne vyhodnotenie výsledkov práce, ktoré sa opiera o meranie a štatistické metódy, prípadne matematické dôkazy. V prípade nameraných hodnôt musí autor opísať metódu merania, priebeh merania, výsledky a interpretáciu výsledkov v kontexte riešeného problému a stanovených cieľov. Na základe vyhodnotenia riešenia autor opíše prínosy svojej práce. Vyhodnocovacia časť tvorí zvyčajne ¼ jadra práce.

4 Záver

Záver práce obsahuje zhrnutie výsledkov práce s jasným opisom prínosov a pôvodných (vlastných) výsledkov autora a vyhodnotenie splnenia stanovených cieľov. Je to stručné zhrnutie informácií uvedených v záverečnej práci. Záver by nemal obsahovať nové informácie.

V závere by mal tiež autor poukázať na prípadné otvorené otázky, ktoré sú nad rámec rozsahu práce a mal by odporučiť ďalšie aktivity na pokračovanie pri riešení problému. Rozsah záveru je minimálne 1 celá strana.

Zoznam príloh

A Karel Language Reference

Karel's Primitives

- `void movek()` - Moves *Karel* one intersection forward.

- `void turn_left()` - Pivots *Karel* 90 degrees left.

- `void pick_beeper()` - Takes a beeper from the current intersection and puts it in the beeper bag.

- `void put_beeper()` - Takes a beeper from the beeper bag and puts it at the current intersection.

- `void turn_on(char* path)` - Turns *Karel* on.

- `void turn_off()` - Turns *Karel* off.

Karel's Sensors

- `int front_is_clear()` - Returns 1 if there is no wall directly in front of *Karel*. `0` if there is.

- `int right_is_clear()` - Returns 1 if there is no wall immediately to *Karel*'s right. `0` if there is.

- `int beepers_present()` - Returns 1 if *Karel* is standing at an intersection that has a beeper. `0` otherwise.

- `int facing_north()` - Returns 1 if *Karel* is facing north. `0` otherwise.

- `int beepers_in_bag()` - Returns 1 if there is at least one beeper in *Karel*'s beeper bag. `0` if the beeper bag is empty.

Misc Functions

- `void set_step_delay(int)` - Sets delay of one *Karel*'s step in miliseconds.

- `loop(int)` - Repeats *Karel*'s instruction in a loop.