Katedra informatiky Přírodovědecká fakulta Univerzita Palackého v Olomouci

BAKALÁŘSKÁ PRÁCE

Demonstrace práce s datovými strukturami



2018

Vedoucí práce: Mgr. Tomáš Kühr, Ph.D.

Patrik Becher

Studijní obor: Aplikovaná informatika, prezenční forma

Bibliografické údaje

Autor: Patrik Becher

Název práce: Demonstrace práce s datovými strukturami

Typ práce: bakalářská práce

Pracoviště: Katedra informatiky, Přírodovědecká fakulta, Univerzita

Palackého v Olomouci

Rok obhajoby: 2018

Studijní obor: Aplikovaná informatika, prezenční forma

Vedoucí práce: Mgr. Tomáš Kühr, Ph.D.

Počet stran: 25

Přílohy: 1 CD/DVD

Jazyk práce: český

Bibliograpic info

Author: Patrik Becher

Title: Data Structure Demonstration

Thesis type: bachelor thesis

Department: Department of Computer Science, Faculty of Science, Pa-

lacký University Olomouc

Year of defense: 2018

Study field: Applied Computer Science, full-time form

Supervisor: Mgr. Tomáš Kühr, Ph.D.

Page count: 25

Supplements: 1 CD/DVD

Thesis language: Czech

Anotace

Cílem bakalářské práce bylo vytvořit nástroj pro podporu výuky algoritmizace, konkrétně práce se základními stromovými datovými strukturami (binární vyhledávací stromy, AVL stromy, červenočerné stromy). Výsledná aplikace podporuje vizualizaci vybraných datových struktur, včetně názorné demonstrace běžně prováděných operací s těmito datovými strukturami se souběžným zobrazením pseudokódu prováděné operace.

Synopsis

Toto mně doplní Míša... :D

Klíčová slova: Binární vyhledávací stromy, Binární strom, AVL strom, Červenočerný strom, Stromové animace Java, JavaFX

Keywords: Binary search trees, Binary Tree, AVL tree, Redblack tree, Tree animations, Java, JavaFX

Rád bych poděkoval panu Mgr. Tomáši Kührovi Ph.D. za vedení této bakalářske práce a panu RNDr. Arnoštu Večerkovi za odbornou pomoc a poskytnuté materiály k práci. Dále bych chtěl poděkoval mé rodině a přítelkyni za podporu přtvorbě.
Místopřísežně prohlašuji, že jsem celou práci včetně příloh vypracoval/a samostatně a za použití pouze zdrojů citovaných v textu práce a uvedených v seznamu literatury.
datum odevzdání práce podpis autora

Obsah

1	Úvo	od .	7			
2	Stromy					
	2.1	Binární strom	10			
	2.2	Binární vyhledávací strom	10			
		2.2.1 Vyhledávání	11			
		2.2.2 Vkládání	13			
		2.2.3 Odebírání	16			
	2.3	AVL strom	18			
Zá	Závěr					
Conclusions						
\mathbf{A}	A První příloha					
В	B Druhá příloha					
\mathbf{C}	C Obsah přiloženého CD/DVD					
Literatura						

Seznam obrázků

1	Příklady neorientovaných stromů	8
2	Popis stromu	
3	Binární strom	
4	Rozdílné výšky stromu	
5	Vyhledávání hodnoty 5	
6	Vkládání hodnoty 6 - hledání	14
7		14
8	·	17
9		19
10	Výpočet faktoru vyvážení pro uzel u	19
11	Ukázka vyváženého AVL stromu	
Sezn	am zdrojových kódů	
1	search	12
2	Result - Třída	
3	search - úprava	15
4	insert	15
5	delete	

1 Úvod

Tato aplikace vznikla za účelem výuky základních binárních stromů. Obsahuje podporu pro Binární vyhledávací, AVL a Červenočerné stromy. Program pomocí animací zobrazuje operace: Vyhledávání, Vkládání a Odebírání prvků ze stromů. Souběžně s animací zobrazuje stručný pseudokód aktuálně prováděné operace. Dále umožňuje Opakovat poslední operaci a Generování náhodných stromů.

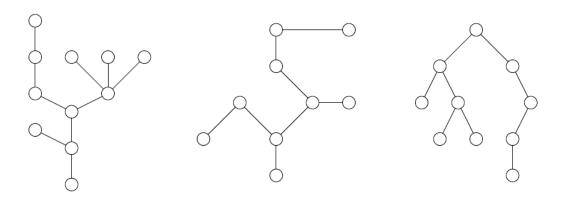
Text samotné práce se dělí na dvě části: *Teoretickou část*, ve které se zabývám teorií vybraných stromů a *Programovou část*, která popisuje samotnou implementaci a funkcionalitu programu.

2 Stromy

V kapitole jsou vysvětleny základní pojmy, které jsou nezbytné k pochopení vlastností stromů obsažených v aplikaci. V podkapitolách jsou osvětleny principy pro tvorbu a následnou práci s konkrétními binárními vyhledávacími stromy. Využitím těchto principů byl naprogramován tento výukový nástroj.

Definice 1 (Strom)

Strom je neorientovaný ¹ souvislý ² graf bez kružnic ³ [1].



Obrázek 1: Příklady neorientovaných stromů

Strom je datová struktura, která představuje stromovou strukturu propojených $uzl\mathring{u}^4$. Uzly jsou mezi sebou vzájemně spojeny pomocí $hran^5$. Strom složený z uzl \mathring{u} U má |U-1| hran.

Definice 2 (Kořenový strom)

Kořenový strom je strom, ve kterém je vybrán jeden vrchol (kořen). Může to být kterýkoliv vrchol. Bývá to ale vrchol, který je v nějakém smyslu na vrcholu hierarchie objektů, která je stromem reprezentována. [1]

¹Mezi každými dvěma vrcholy existuje právě jedna cesta.

 $^{^2}$ Vynecháním libovolné hrany vznikne nesouvislý graf.

³Přidáním jakékoli hrany vznikne graf s kružnicí.

⁴Prvek obsahující hodnotu.

⁵Představuje cestu mezi spojenými uzly.

Důležité pojmy:

- Uzel Jednoduše jakýkoliv prvek stromu.
- **Kořen** Jeden konkrétní uzel, který se nachází na vrcholu stromu. Pouze tento uzel nemá *rodiče*.
- Potomek, následník Uzel, který je hranou přímo připojen k jinému uzlu, cestou od kořene.
- Rodič, předchůdce Uzel, který má alespoň jednoho potomka.
- Sourozenci Skupina uzlů, které mají stejného rodiče.
- **Podstrom** Část stromu, která je úplným stromem, s tím, že kořen tohoto podstromu má svého rodiče.
- Koncový uzel, list Uzel bez potomků.
- Výška stromu Nejdelší délka cesty od kořene k uzlu.



Obrázek 2: Popis stromu

Definice 3 (*m*-ární stromy)

Kořenový strom se nazývá m-ární, právě když každý jeho vrchol má nejvýše m potomků. 2-ární strom se nazývá binární. Kořenový strom se nazývá úplný m-ární, právě když každý jeho vrchol nemá buď žádného nebo má právě m potomků.[2]

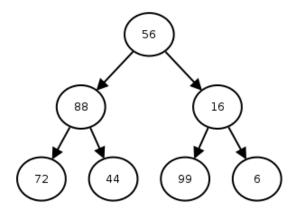
2.1 Binární strom

Definice 4 (Binární strom)

Binární strom je typ kořenových stromů, ve kterém každý obsažený uzel má maximálně 2 potomky.

Každý uzel obsahuje tyto vlastnosti:

- Klíč Hodnota uložená v uzlu.
- Ukazatel na levého potomka
- Ukazatel na pravého potomka
- Ukazatel na jednoho rodiče Tento ukazatel není povinný.



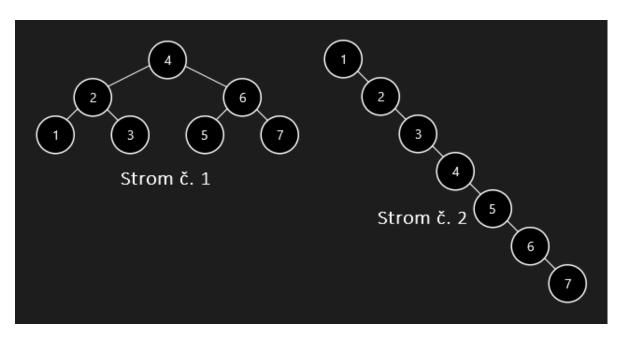
Obrázek 3: Binární strom

2.2 Binární vyhledávací strom

Binární *vyhledávací* strom (zkratka BVS), je speciální typ binárního stromu, kde platí následující:

- Každý pravý potomek P rodiče R má vyšší hodnotu h než jeho rodič. Platí tedy: $P.h > R.h \Rightarrow \text{Pravý}$ podstrom uzlu R obsahuje pouze uzly, které mají vyšší hodnotu než uzel R.
- Každý levý potomek L rodiče R má nižší hodnotu h než jeho rodič. Platí tedy: $P.h > L.h \Rightarrow \text{Lev}$ ý podstrom uzlu R obsahuje pouze uzly, které mají nižší hodnotu než uzel R.
- Ve stromě se nenachází dva uzly se stejnou hodnotou.

Toto uspořádání uzlů v BVS usnadňuje vyhledávání. Operace nad BVS stromem s výškou h mají časovou složitost $\theta(h)$. V nejhorším případě může mít BVS výšku rovnu n-1, kde n je počet uzlů. Oba případy jsou zobrazeny na obrázku 4.



Obrázek 4: Rozdílné výšky stromu

2.2.1 Vyhledávání

Operace vyhledávání patří k nejčastěji používané operaci s BVS. Při vyhledávání je potřeba zadat hodnotu x, kterou chceme vyhledat. Postupně dochází k porovnávání hodnot uzlů s x. Výsledkem vyhledávání je buď uzel, který obsahuje hodnotu x nebo takový uzel neexistuje.

Přesný postup vyhledávání:

1. krok – počáteční

Na začátku vyhledávání je třeba určit aktuální uzel, který označíme u. Hledanou hodnotu označíme x.

V tomto kroku u bude kořen stromu, pokud strom nemá kořen, hodnota x je nenalezena a tím vyhledávání končí.

2. krok – průběžný

Zde dochází k porovnávání x a hodnoty aktuálního uzlu u. Hodnotu u označíme u.h.

Pokud je u prázdný, vyhledávání končí neúspěchem.

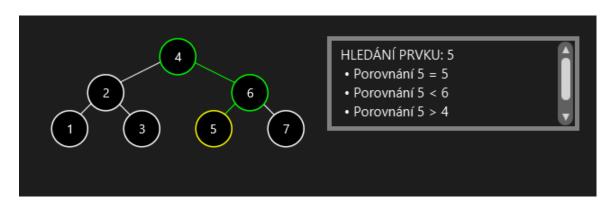
Při porovnávání mohou nastat tyto možnosti:

• x > u.h

V tomto případě jako u nastavíme pravého potomka u. A znovu provedeme 2. krok.

- x < u.hV tomto případě jako u nastavíme levého potomka u. A znovu provedeme 2. krok.
- x = u.hHledaná hodnota x byla nalezena v u. Vyhledávání tedy končí.

Na obrázku 5 je zvýrazněna cesta průchodů stromem při vyhledávání uzlu s hodnotou 5. Na obrázku je i zaznamenána historie porovnávání.



Obrázek 5: Vyhledávání hodnoty 5

Zdrojový kód vyhledávání v jazyku Java:

```
Node search(int value, Node node) { //value je hledaná hodnota, node
       je při prvním volání kořen
    if (node == null) {
      return null; //uzel nebyl nalezen
3
 4
    if (value > node.getValue()) { //pokud je hledaná hodnota vyšší ne
 6
        ž má aktuální uzel
7
      search(value, node.getRight()); //nastavím aktuální uzel na pravé
         ho potomka
     } else if (value < node.getValue()) { //pokud je hledaná hodnota</pre>
8
        vyšší než má aktuální uzel
      search(value, node.getLeft()); //nastavím aktuální uzel na levého
9
          potomka
     } else { //pokud není vyšší ani nižší, tak se musí rovnat
10
      return node; //vrátím nalezený uzel
12
13 }
```

Zdrojový kód 1: search

2.2.2 Vkládání

Při vkládání zadané hodnoty x je nejprve nutné prohledat strom, jestli se zde x již nenachází. Pokud je nalezen uzel s hodnotou x, je výsledkem operace nalezený uzel. V případě, že daný strom tento uzel neobsahuje, je jako potomek posledního prohledávaného uzlu vložen nový uzel s hodnotou x.

Přesný postup vkládání:

1. krok – počáteční

Na začátku vkládání je třeba určit aktuální uzel, který označíme u. Vkládanou hodnotu označíme x.

V tomto kroku u bude kořen stromu. Pokud strom nemá kořen, vytvoříme nový uzel s hodnotou x a ten vložíme do stromu. Uzel se stane kořenem stromu, čímž vkládání končí.

2. krok – vyhledávání, vkládání

Před vkládáním je nejprve nutno ověřit, zde se ve stromu nenachází uzel s hodnotou x.

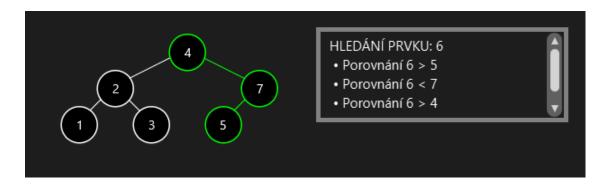
Vyhledávání může dopadnout těmito způsoby:

- \bullet Prvek byl nalezen. Pokud byl uzel s hodnotou x nalezen, vkládání končí neúspěchem.
- Prvek nebyl nalezen, přičemž platí $x > u.h^6$. Vytvoříme nový uzel s hodnotou x a vložíme jako pravého potomka u^7 .
- Prvek nebyl nalezen, přičemž platí $x < u.h^6$. Vytvoříme nový uzel s hodnotou x a vložíme jako pravého potomka u^7

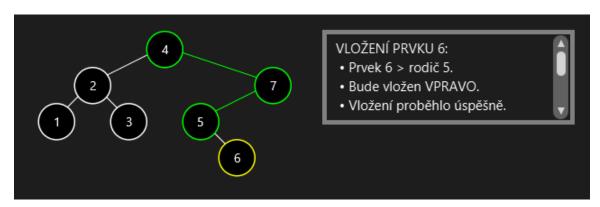
Na obrázku 6 je zvýrazněna cesta průchodů stromem při vyhledávání uzlu s hodnotou 6. Samotné vložení je na obrázku 7.

⁶Hodnota posledního navštíveného uzlu při hledání.

⁷Poslední navštívený uzel při hledání.



Obrázek 6: Vkládání hodnoty 6 - hledání



Obrázek 7: Vkládání hodnoty 6 - vložení

Zdrojový kód vkládání v jazyku Java:

Před samotnou funkcí insert je třeba vytvořit třídu Result, která bude dále použita:

```
1 Class Result() {
2  private Node node; //poslední navštívený uzel (hledaný/rodič)
3  private boolean isFind; //parametr pro určení zda byl uzel nalezen
4  public Result(Node node, boolean isFind) { //konstruktor
6  ...
7  }
8  ...
9 }
```

Zdrojový kód 2: Result - Třída

Dále je potřeba trochu poupravit již známou funkci search 1:

```
1 Result search(int value, Node node) { //value je hledaná hodnota,
      node je při prvním volání kořen
    if (value > node.getValue()) {
      if (node.getRight() != null) {
4
       search(value, node.getRight()); //pokud má pravého potomka
5
     } else {
       return new Result(node, false); //pokud nemá pravého potomka,
 6
          node = rodič vkládaného
7
     }
8
    } else if (value < node.getValue()) {</pre>
      if (node.getLeft() != null) {
9
10
       search(value, node.getLeft()); //pokud má levého potomka
11
      } else {
       return new Result(node, false); //pokud nemá levého potomka,
           node = rodič vkládaného
13
     }
   } else {
     return new Result(node, true);
16
17 }
```

Zdrojový kód 3: search - úprava

Nyní funkce insert:

```
1 Node insert(int value) {
    Result result = search(value); //nejprve vyhledáme value
    if (result.isFind()) { //pokud byl uzel s danou hodnotou nalezen
     return result.getNode();
5
    } else { //dále vkládáme nově vytvořený uzel:
     if (value > result.getNode().getValue()) {
6
7
       result.getNode().setRight(new Node(value));//bude pravý potomek
      } else {
       result.getNode().setLeft(new Node(value));//bude levý potomek
10
    }
11
    return null;
12
```

Zdrojový kód 4: insert

2.2.3 Odebírání

Při odebírání zadané hodnoty x je stejně jako u vkládání nejprve nutné prohledat strom, zda se odebíraný uzel s hodnotou x ve stromě nachází. Pokud je uzel nalezen, je následně smazán a struktura stromu případně upravena.

Přesný postup odebírání:

1. krok – počáteční

Na začátku odebírání je třeba určit aktuální uzel, který označíme u. Odebíranou hodnotu označíme x.

V tomto kroku u bude kořen stromu. Pokud strom nemá kořen odebírání končí, uzel s hodnotou x nebyl nalezen.

2. krok – vyhledávání

Před odebíráním je nejprve nutno uzel s hodnotou x vyhledat.

Vyhledávání může dopadnout těmito způsoby:

- \bullet Prvek nebyl nalezen. Pokud se uzel s hodnotou x ve stromě nenachází, odebírání končí neúspěchem.
- Prvek byl nalezen.
 Nyní můžeme nalezený uzel u odebrat.

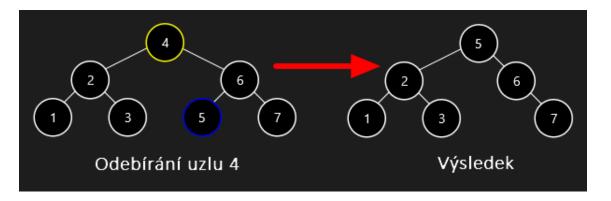
3. krok – odebírání

Odebírání u má tyto možnosti:

- Pokud u je list⁸.
 List u může být odebrán.
- u má jednoho potomka.
 u bude nahrazen podstromem potomka.
- u má dva potomky.
 - u bude nahrazen nejlevějším prvkem z pravého podstromu.
 - -u bude nahrazen nejpravějším prvkem z levého podstromu.

V následujícím obrázku 8 je ukázka mazání uzlu s hodnotou 4, který je následně nahrazen uzlem s hodnotou 5, který je nejlevější prvek z pravého podstromu.

⁸Nemá žádné potomky.



Obrázek 8: Odebírání hodnoty $4\,$

Zdrojový kód odebírání v jazyku Java:

 $^{^9\}mathrm{Tento}$ kód je pouze zjednodušená implementace k lepšímu pochopení odebírání.

```
1 boolean delete(int value) {
    Node removeNode = result.getNode();
    Node helpNode; //pomocná proměnná
    Result result = search(value); //vyhledám hodnotu
 6
    if (!result.isFind()) { //pokud ho nenajdu
8
      return false;
9
10
     if ((removedNode.getLeft() != null) && (removedNode.getRight() !=
11
        null)) { //Pokud má dva potomky
12
      helpNode = removeNode.getRight(); //dosadím pravého potomka
13
      while (helpNode.getLeft() != null) { //a hledám nejlevějšího
14
15
       helpNode = helpNode.getLeft();
16
17
      removeNode.setNode(helpNode);
18
     } else if (removedNode.getLeft() != null) { //pouze levý potomek
20
     removeNode.setNode(removedNode.getLeft());
     } else if (removedNode.getRight() != null) { //pouze pravý potomek
2.1
22
     removeNode.setNode(removedNode.getRight());
    } else { //pokud je to list
      removeNode.delete(); //odstraním list ze stromu
2.4
2.5
26
27
    return true;
28 }
```

Zdrojový kód 5: delete

2.3 AVL strom

Určitým způsobem by se dalo tvrdit, že *AVL strom* je binární vyhledávací strom, který má ale jednu zásadní odlišnost. Složitost všech operací u BVS je závislá na výšce stromu, je tedy žádoucí udržovat výšku stromu co nejnižší. *AVL strom* je tedy vyvážený binární vyhledávací strom.

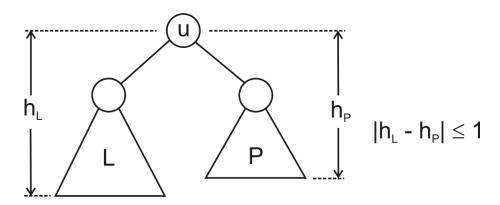
Definice 5 (Vyvážený strom)

Strom je vyvážený tehdy a jen tehdy, je-li rozdíl výšek každého uzlu nejvýše 1.[5][3]

V následujícím obrázku 11 je předchozí definice o vyváženém stromu názorně vysvětlena. Nejprve je třeba zavést pojmy:

- u aktuální uzel.
- L levý podstrom u.

- P pravý podstrom u.
- h_L výška levého podstromu.
- h_P výška pravého podstrom.

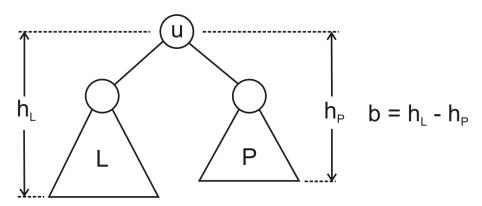


Obrázek 9: Vyvážený strom

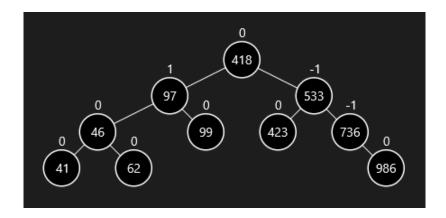
Faktor vyvážení

Pro kontrolu dodržení pravidla o vyvážení, je třeba pro každý uzel zavést novou vlastnost faktor vyvážení uzlu. Tuto vlastnost budeme značit b (anglicky balance). b obsahuje informaci o aktuálním vyvážení daného podstromu.

Nabývá hodnot <-2, 2>, přičemž uzel je vyvážený pokud jeho faktor b je 1, 0 nebo -1. Při operaci přidávání nebo odebírání, může b být 2 nebo -2, pak je zapotřebí provést transformaci, která dosáhne vyvážení daného uzlu.



Obrázek 10: Výpočet faktoru vyvážení pro uzelu



Obrázek 11: Ukázka vyváženého AVL stromu

Každý uzel ALV stromu obsahuje tyto vlastnosti:

- $\bullet~\mathbf{Klíč}$ Hodnota uložená v uzlu.
- Faktor vyvážení Faktor vyvážení daného uzlu.
- Ukazatel na levého potomka
- Ukazatel na pravého potomka
- Ukazatel na jednoho rodiče Tento ukazatel není povinný.

Závěr

Závěr práce v "českém" jazyce.

Conclusions

Thesis conclusions in "English".

A První příloha

Text první přílohy

B Druhá příloha

Text druhé přílohy

C Obsah přiloženého CD/DVD

Na samotném konci textu práce je uveden stručný popis obsahu přiloženého CD/DVD, tj. jeho závazné adresářové struktury, důležitých souborů apod.

bin/

Instalátor Instalator programu, popř. program Program, spustitelné přímo z CD/DVD. / Kompletní adresářová struktura webové aplikace Webovka (v ZIP archivu) pro zkopírování na webový server. Adresář obsahuje i všechny runtime knihovny a další soubory potřebné pro bezproblémový běh instalátoru a programu z CD/DVD / pro bezproblémový provoz webové aplikace na webovém serveru.

doc/

Text práce ve formátu PDF, vytvořený s použitím závazného stylu KI PřF UP v Olomouci pro závěrečné práce, včetně všech příloh, a všechny soubory potřebné pro bezproblémové vygenerování PDF dokumentu textu (v ZIP archivu), tj. zdrojový text textu, vložené obrázky, apod.

src/

Kompletní zdrojové texty programu Program / webové aplikace Webovka se všemi potřebnými (příp. převzatými) zdrojovými texty, knihovnami a dalšími soubory potřebnými pro bezproblémové vytvoření spustitelných verzí programu / adresářové struktury pro zkopírování na webový server.

readme.txt

Instrukce pro instalaci a spuštění programu PROGRAM, včetně všech požadavků pro jeho bezproblémový provoz. / Instrukce pro nasazení webové aplikace WEBOVKA na webový server, včetně všech požadavků pro její bezproblémový provoz, a webová adresa, na které je aplikace nasazena pro účel testování při tvorbě posudků práce a pro účel obhajoby práce.

Navíc CD/DVD obsahuje:

data/

Ukázková a testovací data použitá v práci a pro potřeby testování práce při tvorbě posudků a obhajoby práce.

install/

Instalátory aplikací, runtime knihoven a jiných souborů potřebných pro provoz programu Program / webové aplikace Webovka, které nejsou standardní součástí operačního systému určeného pro běh programu / provoz webové aplikace.

literature/

Vybrané položky bibliografie, příp. jiná užitečná literatura vztahující se k práci.

U veškerých cizích převzatých materiálů obsažených na CD/DVD jejich zahrnutí dovolují podmínky pro jejich šíření nebo přiložený souhlas držitele copyrightu. Pro všechny použité (a citované) materiály, u kterých toto není splněno a nejsou tak obsaženy na CD/DVD, je uveden jejich zdroj (např. webová adresa) v bibliografii nebo textu práce nebo v souboru readme.txt.

Literatura

- [1] BĚLOHLÁVEK, Radim. Algoritmická matematika 2 část 1 [online]. 2012-05-15. [cit. 2018-07-07]. Dostupné z: http://belohlavek.inf.upol.cz/vyuka/algoritmicka-matematika-2-1.pdf
- [2] BĚLOHLÁVEK, Radim; VYCHODIL, Vilém. Diskrétní matematika pro informatiky II [online]. 2010-10-16. [cit. 2018-07-07]. Dostupné z: http://belohlavek.inf.upol.cz/vyuka/dm2.pdf
- [3] DVORSKÝ, Jiří. Algoritmy I.[online]. 2007-02-27. [cit. 2018-07-07]. Dostupné z: http://www.cs.vsb.cz/dvorsky/Download/SkriptaAlgoritmy/Algoritmy.pdf
- [4] BINARY SEARCH TREES Dostupné z: http://cs.umw.edu/finlayson/class/fall12/cpsc230/notbinary-search-trees.html
- [5] ADELSON-VELSKII, G. M.; LANDIS, Ε. Μ. algorithm An ganization for orof information. Soviet Mathematics Doklady, 1962. [online]. 2018-07-07] 3:1259-1263, [cit. Dostupné http://professor.ufabc.edu.br/ jesus.mena/courses/mc3305-2q-2015/AED2-10-avl-paper.pdf