



UNIUNEA EUROPEANĂ



GUVERNUL ROMÂNIEI
MINISTERUL MUNCII, FAMILIEI
ȘI PROTECȚIEI SOCIALE
AMPOSDRU



Fondul Social European
POSDRU 2007-2013



Instrumente Structurale
2007-2013



MINISTERUL
EDUCAȚIEI
CERCETĂRII
TINERETULUI
ȘI SPORTULUI

OIPOSDRU



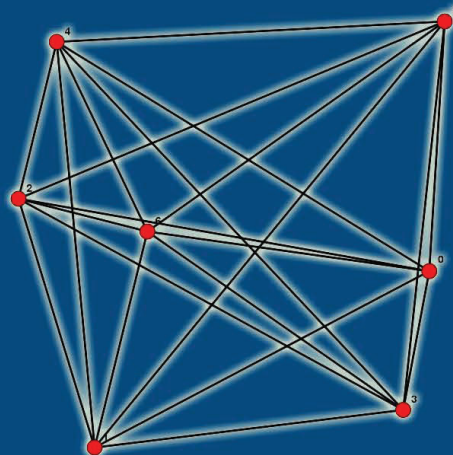
UNIVERSITATEA
POLITEHNICĂ DIN
BUCUREȘTI

Proiect cofinanțat din Fondul Social European prin Programul Operațional Sectorial Dezvoltarea Resurselor Umane 2007-2013
Investește în oameni!

Răzvan ANDONIE Angel CAȚARON Zoltán GÁSPÁR Honorius GÂLMEANU
Mihai IVANOVICI István LŐRENTZ Lucian SASU

ALGORITMI , ȘI STRUCTURI DE DATE

Aplicații în Imagistică și Bioinformatică



Editura Universității *TRANSILVANIA* din Brașov

© 2012 EDITURA UNIVERSITĂȚII *TRANSILVANIA* DIN BRAȘOV

Adresa: 500091 Brașov,
B-dul Iuliu Maniu 41A
Tel:0268 – 476050
Fax: 0268 476051
E-mail : editura@unitbv.ro



Tipărit la:
Tipografia Universității *Transilvania* din Brașov
B-dul Iuliu Maniu 41A
Tel: 0268 – 476050

Toate drepturile rezervate

Editură acreditată de CNCSIS
Adresa nr.1615 din 29 mai 2002

Referenți științifici: Prof. univ. dr. ing. Gheorghe TOACȘE
Prof. univ. dr. ing. Vasile BUZULOIU

Descrierea CIP a Bibliotecii Naționale a României

Algoritmi și structuri de date : aplicații în imagistică și bioinformatică /

Răzvan Andonie, Angel Cațaron, Zoltan Gáspár, - Brașov : Editura
Universității "Transilvania", 2012

Bibliogr.

ISBN 978-606-19-0058-9

I. Andonie, Răzvan

II. Cațaron, Angel

III. Zoltan, Gáspár

004.421(075.8)

510.5(075.8)

004.422.63(075.8)

Cuprins

1	Structuri de date generice	7
1.1	Breviar teoretic	7
1.2	Java Collections Framework	10
1.2.1	Liste în Java: clasa <i>ArrayList</i>	10
1.2.2	Mulțimi în Java: clasele <i>HashSet</i> și <i>TreeSet</i>	12
1.3	C++ Standard Template Library	13
2	Grafuri în Java	17
2.1	Breviar teoretic	17
2.2	Izomorfismul grafurilor	18
2.3	Pachetul de clase JGraphT	20
2.4	Folosirea grafurilor în biochimie	21
3	Arbori în C++	25
3.1	Breviar teoretic	25
3.2	Supertree	26
3.3	Determinarea supertree-ului unor subarbori	30
4	Programare paralelă în CUDA	33
4.1	Breviar teoretic	33
4.2	CUDA	38
4.2.1	Terminologie, abrevieri	40
4.2.2	Arhitectura CUDA	40
4.2.3	Modelul de programare	41
4.2.4	Lansarea în execuție a threadurilor	44
4.2.5	Compilarea exemplelor	45
4.3	Primul program CUDA	46
4.4	Adunarea a doi vectori	48
4.5	Însumarea elementelor unui vector	50
4.5.1	Accesul la memoria globală	51
4.5.2	Bariera de sincronizare	51
4.6	Înmulțirea matricelor	52
4.6.1	Matrice	53
4.6.2	Cazul înmulțirii secvențiale	53

4.6.3	Cazul paralel	53
5	Aplicații în CUDA	57
5.1	Texturi și imagini color	57
5.1.1	Procesarea texturilor	57
5.2	Interoperabilitatea cu OpenGL	60
5.3	Folosirea bibliotecii CUDA FFT 2D	64
5.4	Programarea generică folosind Thrust	67
5.4.1	Transformarea generică	68
5.5	Alinierea secvențelor folosind CUDA	69
5.5.1	Algoritmul Needleman-Wunsch de aliniere globală	70
6	Algoritmi de aproximare	77
6.1	Breviar teoretic	77
6.2	Introducere în Python	78
6.3	Problema celui mai scurt superstring	81
7	Algoritmi euristici în grafuri	85
7.1	Breviar teoretic	85
7.1.1	Grafuri în Python cu <code>igraph</code>	87
7.2	Înălțimea unui nod într-o rețea filogenetică	91
7.3	Graph matching	92
A	Surse aferente capitolului 4	97
B	Surse aferente capitolului 5	107

Prefață

Lucrarea de față a fost realizată în perioada septembrie 2010 - martie 2012, în cadrul proiectului POS-DRU/86/1.2/S/61756 intitulat *Tehnici de Analiză, Modelare și Simulare pentru Imagistică, Bioinformatică și Sisteme Complexe (ITEMS)*. Colectivul ITEMS de la Universitatea Transilvania din Brașov a răspuns de conceperea și predarea cursului de *Algoritmi și Structuri de Date* pentru programul de Masterat de Excelență organizat sub umbrela acestui proiect la Universitatea Politehnica din București. Prezentul volum conține lucrările de laborator ale acestui curs. Chiar dacă a fost conceput pentru masteranzii ITEMS, suntem convinși că lucrarea poate fi utilă multor studenți și specialiști din domeniul Calculatoare și Tehnologia Informației.

Încă de la început, ne-am propus să concepem cursul altfel decât este el predat la majoritatea universităților. Structurile de date și algoritmi sunt desigur standard, deoarece este practic imposibilă o schimbare radicală a acestui domeniu. Totuși, ne-am concentrat mai ales pe tehnici specializate, mai rar prezentate. Și atunci ce este de fapt nou, în afară de aplicațiile alese și implementate de noi?

Nouă este integrarea a trei concepte. Primul dintre acestea este filozofia noastră de predare, care se bazează pe structuri de date generice orientate pe obiect. Am folosit Java, C++ și Python, toate limbaje care permit o astfel de abordare, pentru a sublinia că nu contează limbajul de programare, ci modul de abordare. C++ este alegerea naturală pentru programarea paralelă în CUDA. Python este frecvent utilizat pentru programare *ad-hoc* sau prototipare rapidă, fiind răspândit mai ales în grupurile de cercetare. În sfârșit, Java se impune prin portabilitate, fiind totodată, alături de C/C++, un limbaj de referință. Al doilea concept este că viitorul aparține programării paralele. Din această cauză, o parte din lucrările de laborator sunt implementate în CUDA, pe plăci grafice cu arhitectură masiv paralelă. Al treilea concept este un accent mare pus pe teoria complexității, atât în cazul secvențial, cât și în cazul paralel. Analiza complexității algoritmilor folosiți este un pas important, chiar și atunci când ne referim la metodele definite în containere generice.

Lucrările de laborator prezentate au fost selectate pentru a ilustra aplicații în imagistică și bioinformatică. Fișierele și programele aferente pot fi accesate prin Internet, la adresa: <http://miv.unitbv.ro/asd>.

Echipa de autori este foarte omogenă. Ne cunoaștem de ani buni și am realizat multe lucrări de cercetare împreună. Practic, putem vorbi despre o școală de algoritmi la Brașov. O parte dintre noi suntem cadre didactice universitare, o alta lucrăm în firme de software. Pentru noi, perioada elaborării acestui material a fost extrem de creativă. Ceea ce a determinat acest lucru a fost în special lucrul în echipă. Este motivul pentru care autorii sunt listați alfabetic, fiecare cu contribuții egale, greu de separat.

Brașov, martie 2012.

Autori