

Experimentální analýza algoritmů je empirickým protějškem teoretické analýzy algoritmů.

Je součástí širší disciplíny, která se nazývá **Experimentální algoritmika (EA)**.

EA vznikla někdy po roce 1990, její vznik a rozvoj byl umožněn možnostmi moderních počítačů, zejména:

Provádět velké množství experimentů s rozsáhlými daty

Elektronicky publikovat výsledky

Vytvářet a sdílet knihovny efektivních implementací algoritmů

Vytvářet a sdílet soubory reálných i simulovaných testovacích dat

Cíle EA:

1) Uvádět algoritmy navržené teoretiky do praxe – to obnáší zejména:

Implementace, ladění, testování správnosti

Měření charakteristik (např. času, počtu určitých operací, výpadků z cache atd.)

Vytváření programových systémů jako vývojového prostředí pro tyto činnosti

Vytváření knihoven efektivních programů (včetně dokumentace)

Generování a vytváření knihoven testovacích dat

Tvorba prostředků pro vizualizaci, animaci, prezentaci výsledků

2) Vyhodnocením výsledků experimentů zpětně napomáhat teoretikům při navrhování lepších algoritmů – např.:

Porovnávat algoritmy se stejnou asymptotickou složitostí

Vyčíslovat „skryté“ konstanty v symbolech O , Ω , θ

Suplovat teoretické výsledky v případech, kdy nejsou známy

Studovat chování heuristických metod

Studovat vliv hierarchických pamětí na chování algoritmů

Formulovat nové otázky a hypotézy pro teoretickou analýzu

Následující odkazy mohou být užitečné, ale není zaručeno, že v nich najdete všechno, co budete pro svou práci potřebovat. Některé také během času už nemusí být aktuální, na druhé straně se neustále objevují další.

Úvodní a přehledové články o EA:

B. M. E. Moret: Towards a discipline of Experimental algorithmics, in DIMACS series in Discrete Math. and Theoretical Comp. Sci. 59 (2002), 197 - 213 (původně in Proc. 5th DIMACS Challenge Workshop, 1996)

http://lcbp.epfl.ch/~moret/dimacs_algorithmics.pdf

Obsah: motivace, předmět zájmu a cíle EA

C. C. McGeoch: A bibliography of algorithm experimentation, tamtéž

Obsah: komentovaný přehled užitečné literatury z různých oborů

C. Demetrescu, G. F. Italiano: What do we learn from Experimental algorithmics?, in Proc. MFCS 2000, LNCS 1893, 36 – 51

<http://www.springerlink.com/content/yj370kk82q958drx/fulltext.pdf>

Obsah: přehled existujících softwarových knihoven, souborů testovacích dat, systémů

Softwarové knihovny:

LEDA (Library of Efficient Data Types and Algorithms)

<http://www.algorithmic-solutions.com/leda/>

původní autoři K. Mehlhorn a S. Näher, 1989, spojeno s knihou

K. Mehlhorn, S. Näher: LEDA – A Platform for Combinatorial and Geometric

Computing, Cambridge Univ. Press, 1999

<http://www.mpi-inf.mpg.de/~mehlhorn/LEDAbook.html>

Stony Brook Algorithm Repository

<http://www.cs.sunysb.edu/~algorithm/>

spojeno s knihou

S. Skiena: Algorithm Design Manual, Springer 1997, 2. vydání 2008

<http://www.springer.com/computer/theoretical+computer+science/foundations+of+computations/book/978-1-84800-069-8>

Knihovny testovacích dat:

Stanford GraphBase

Spojena s D. Knuthem a jeho knihou The Stanford GraphBase: A Platform for Combinatorial Computing, ACM Press, Addison-Wesley 1993 (paperback 2009)

<http://www-cs-faculty.stanford.edu/~knuth/sgb.html>

Vyplatí se navštívit i osobní stránky D. Knutha

<http://www-cs-faculty.stanford.edu/~knuth/index.html>

Knihovna CATS (Combinatorial Algorithms Test Sets)

Projekt představili A. V. Goldberg a B. M. E. Moret v roce 1999

Spojená s časopisem ACM Journal on Experimental Algorithmics (JEA)

<http://www.jea.acm.org>

<http://www.jea.acm.org/CATS> - stránka knihovny v současné době nefunguje

Nástroje pro vizualizaci a animaci

LEONARDO

<http://www.dis.uniroma1.it/~demetres/Leonardo/>

Komplexní výpočetní prostředí

systém LINK

vyvinut v Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) na Rutgers University

<http://www.dimacs.rutgers.edu/~berryj/LINK.html>

Časopisy věnované EA:

ACM Journal on Experimental Algorithmics (JEA)

<http://www.jea.acm.org>

Konference:

SEA – Symposium on Experimental Algorithm

WEA – Workshop on Experimental Algorithm

ExpCS – Workshop on Experimental Computer Science

ALENEX – Algorithm Engineering & Experiments

Dagstuhl Seminar on Experimental Algorithmics

WAE – Workshop on Algorithm Engineering

Statistický software:

K EA patří i statistická analýza experimentálních výsledků.

Přehled statistického softwaru dostupného na MFF je na stránkách KPMS

<http://www.karlin.mff.cuni.cz/~kpms/?vyber=software&lang=>

Komentovaný přehled nejen softwaru je na

<http://www.stahroun.me.cz/ssysel/>

Náhodné generátory:

Přehled literatury o náhodném generování, softwaru a testů je např. na

<http://random.mat.sbg.ac.at/>

Pro experimentování se používají simulační generátory, nikoli kryptografické

Protředky pro přesné měření času v experimentech:

Čas běhu programu je silně ovlivněn konkrétním výpočetním prostředím (procesor, paměť, operační systém, ...) a způsobem měření času (počítání cyklů procesoru, různé funkce odečítající čas, profily, ...). Je proto třeba hledat prostředky pro konkrétní konfiguraci počítače, operační systém a programovací jazyk. Zde si budete muset při vyhledávání informací poradit každý sám, protože je toho mnoho a navíc se to rychle mění.

Jednoduchá obecně platná doporučení:

Zabránit přístupu jiných uživatelů (odpojit počítač od sítě) a nechat běžet jen testovaný program

Omezit systémové služby na nutné minimum

Opakovat experiment několikrát s naprosto stejnými daty a z naměřených časů vzít např. průměr

Statistické metody pro vyhodnocování experimentálních dat:

Dostupná a srozumitelná literatura je např.:

J. Anděl: Statistické metody, MATFYZPRESS, Praha 1998

K. Zvára, J. Štěpán: Pravděpodobnost a matematická statistika, MATFYZPRESS, Praha 1997

V. Dupač, M. Hušková: Pravděpodobnost a matematická statistika, Karolinum, Praha 1999

J. Antoch, D. Vorlíčková: Vybrané metody matematické statistiky, Academia, Praha 1992

Poněkud starší a obtížnější je:

C. R. Rao: Lineární modely statistické indukce a jejich aplikace, Academia, Praha 1978

Zajímavé je také nahlédnout do statistických tabulek (i když v době existujícího statistického softwaru se bez nich obejdeme), zejména legendárních:

J. Janko: Statistické tabulky, Nakladatelství ČSAV, Praha 1958,

které obsahují kromě samotných tabulek i obsáhlé vysvětlující texty a příklady použití