University of BRISTOL

Bristol Algorithms Days – BAD '09

University of Bristol

11-12 May 09

http://www.cs.bris.ac.uk/Research/Algorithms/bad09.html

Department of Computer Science
University of Bristol, Bristol, UK

EPSRC

bridging the gaps

University of Bedfordshire

Acknowledgements

Organising an event (however small) is a job that involves a lot of hard work. This event is no exception. There have been many people involved and we would like to take this opportunity to thank them.

We owe great thanks to Kerrie Walker and Jean Vogel-Gourgand for their invaluable help in organising all conference facilities, accommodation and for making your experience as enjoyable as possible!

Many thanks also to a team of people including Leon Atkins, Benjamin Sach and Ashley Montanaro who have provided invaluable help and and advice throughout.

We would like to thank all of our speakers for providing their abstracts so promptly, and all the delegates attending BAD ‘09 who will make this event interesting and fun, and help to bridge some of the gaps between Computer Science, Engineering and Mathematics.

Raphaël Clifford and Paul Sant
(Bristol Algorithms Days 2009 organisers)

Bristol Algorithms Days 09 Event

The third annual Bristol Algorithms Days (BAD'09) will be held at the Department of Computer Science, University of Bristol on the 11th and 12th May 2009. It is expected that most participants will arrive on the evening of the 10th, and leave in the evening of the 12th. The aim of BAD'09 is to gather researchers and students interested in the study of theoretical and practical aspects of algorithms.

This workshop is funded by "Bridging the Gaps", an EPSRC-funded initiative to set up an environment that nurtures interdisciplinary working, with the aim of developing new collaborative research programmes at the interface between mathematical sciences, information and computing technology and engineering.

In the spirit of the workshop we invite participants to present work and ideas in many areas of algorithmics including the following:

Combinatorial Pattern Matching
Complexity Theory
Approximation Algorithms
Randomised Algorithms
Graph Theory
Probabilistic Analysis
Game Theory
Quantum Computing

Bridging the gaps

Bridging the Gaps is an EPSRC-funded initiative that involves the Faculty of Engineering and the Department of Mathematics in the Science Faculty at the University of Bristol. £400k has been awarded to Bristol in a competitive call to set up an internal environment that nurtures interdisciplinary working, with the aim of developing new collaborative research programmes at the interface between mathematical sciences, information and computing technology, and engineering. Under the grant, which will run for three years from 2007, a small team comprising Alan Champneys (Eng Maths), Raphaël Clifford (CS), Yves Tournigny (Maths) and Chris Allen (Aero) supported by project management, solicits proposals from Bristol investigators for seminar programmes, away days, workshops, retreats or any imaginative event aimed at bridging the gaps between disciplines. Four broad themes have been identified, being A. Uncertain Systems, B. Pure Mathematics and Algorithms, C. Media and Materials and D. Computational Science, although any genuinely interdisciplinary collaboration that bridges the gaps will be considered. Involvement of researchers from outside the University is a key part of the programme, and success will be judged by the number of new research outputs (papers, grants, research networks) that result. Bristol Algorithms Days 2009 is funded and supported by the Bristol Bridging the Gaps project.

For more information, visit http://bridgingthegaps.bristol.ac.uk/

Participation

The following is the final list of participants for BAD'09:

Invited speakers:
Artur Czumaj
Leszek A. Gasieniec
David Manlove
Harald Raecke
Iain Stewart
Kostas Tsichlas

Contributed speakers:
Martyn Amos
Peter Biro
Raphaël Clifford
James Cruise
Sharon Curtis
James Gate
Alan Gibbons
Andrew Potter
Ida Pu
Pim Van 't Hof
Alexander Tiskin
Andy Twigg
Standa Zivny

Other participants:
Leon Atkins
Manolis Christodoulakis
Ilias Flaounas
Aram Harrow
Rob Irving
Vasileios Lampos
Inbok Lee
Jose Martinez-Carranza
Ashley Montanaro
Alexandru Popa
Benjamin Sach
Paul Sant
Tristan Snowsill
Siva Subramani

Optimising paired and pooled kidney exchanges

David Manlove
Department of Computing Science
University of Glasgow

Recent changes in legislation in the UK have allowed a patient who requires a kidney transplant, and who has a willing but incompatible donor, to be involved in a "pairwise kidney exchange" with another patient-donor pair in a similar position. "Pooled donations" extend this concept to three couples in a cyclic manner. The Organ Donation and Transplantation Directorate of NHS Blood and Transplant run the National Matching Scheme for Paired Donation, which finds pairwise and 3-way exchanges (the latter being pooled exchanges involving three couples) at regular intervals. This matching scheme is based on optimising firstly the number of transplants, and subject to this, the total weight of the transplants, based on a scoring system that incorporates a number of factors including sensitivity, HLA compatibility, age and geographic location.

In this talk we describe a range of computational techniques that we have used in order to construct optimal exchanges for NHS Blood and Transplant on various occasions between April 2008 and April 2009. The first two of these involve polynomial-time algorithms, based on weighted matching in graphs, to find an optimal set of exchanges and are implemented in C++. The third technique deals with an NP-hard optimisation problem, and uses integer linear programming, implemented in Matlab. We present some computational results based on simulations involving random samples from a large dataset of patient-donor pairs, and in addition we present a summary of the results that we obtained for the "real" matching runs, as applied to anonymous data supplied by NHS Blood and Transplant.

This is joint work with Péter Biró and Kirstin MacDonald.

Efficient rumor spreading (gossiping) methods in wireless/radio networks

Leszek A. Gasieniec
University of Liverpool

A communication network is very often modeled as a graph of connections in which the nodes exchange information (messages) via directed or undirected links. An associated communication protocol determines the way the messages are exchanged. Among the most popular network models are: (1) the *message passing model* in which a node in one round can inform all its neighbors; (2) the *telephone model* also known as the *matching model* where in each round edges along which the exchange of messages is performed form a matching in the graph of connections. More recently, due to arrival of wireless technology (3) the *radio network model* attracted more attention in algorithms community. In this model, similarly as in the message-passing model, a message transmitted by a node is destined for all neighbors of this node. However, it is assumed that due to interference a node can successfully receive a message if and only if exactly one of its neighbors transmits during this round.

The two classical problems of disseminating information in computer networks are *broadcasting* (one-to-all communication) and *gossiping* (total information exchange). In broadcasting, the goal is to distribute a piece of information (*broadcast message*) from a distinguished source node to all other nodes in the network. In gossiping, however, each node in the network is expected to distribute its own message to every other node in the network. A lot of attention has been given to the broadcasting problem that resulted in a large volume of efficient algorithmic solutions in the models described above. Surprisingly much less is known about gossiping. The latter problem is more advanced algorithmically (in principle it is a simultaneous multiple broadcasting) thus it must involve more complex communication strategies. Further study on efficient gossiping methods gained recently an extra motivation through an increasing interest in, e.g., information aggregation methods that propel fundamental applications in sensor networks.

During this survey talk we present the most important developments in efficient radio gossiping. We discuss deterministic as well as randomized methods and we consider various models in relation to knowledge and stability of network connections. Finally we state a number of fundamental graph and communication problems in relation to gossiping that require further studies.

Deterministic Structures over P2P Networks

Kostas Tsichlas

Various indexing schemes for P2P environments have been proposed for the manipulation of data. These schemes have become quite popular due to the explosion in use of such networks. We design a new P2P data structure, called the Deterministic Distributed tree (DDtree). The DDtree compares favourably to other designs for the following reasons: a) it divides the overlay structure of the P2P environment from the actual elements stored in it and b) it provides better complexities (which are deterministic) compared to all previous solutions. Additionally, the division between elements and nodes results in a load balancing problem in which we have provided an innovative and very efficient solution. This load-balancing scheme can also be applied to any other tree structure in a P2P environment. Finally, a small discussion on models of P2P Networks is initiated.

Joint work with S. Sioutas and G.S. Brodal.

Oblivious Interference Scheduling

Harald Raecke
University of Warwick

In the *interference scheduling problem*, one is given a set of *n* communication requests described by pairs of points from a metric space. The points correspond to devices in a wireless network. In the *directed version* of the problem, each pair of points consists of a dedicated sending and a dedicated receiving device. In the *bidirectional version* the devices within a pair shall be able to exchange signals in both directions. In both versions, each pair must be assigned a power level and a color such that the pairs in each color class can communicate simultaneously at the specified power levels. The feasibility of simultaneous communication within a color class is defined in terms of the Signal to Interference Plus Noise Ratio (SINR) that compares the strength of a signal at a receiver to the sum of the strengths of other signals. This is commonly referred to as the “physical model” and is the established way of modelling interference in the engineering community. The objective is to minimize the number of colors corresponding to the time needed to schedule all requests.
We study *oblivious power assignments* in which the power value of a pair only depends on the distance between the points of this pair. We prove that oblivious power assignments cannot yield approximation ratios better than (n) for the directed version of the problem. For the bidirectional version we can show the existence of a universally good oblivious power assignment: For any set of *n* bidirectional communication requests, the so-called “square root assignment” admits a coloring with at most polylog(n) times the minimal number of colors.

Joint work with Alexander Fanghänel, Thomas Kesselheim, and Berthold Vöcking.

The one-to-many node-disjoint paths problem in certain interconnection networks

Iain Stewart
University of Durham

The one-to-many node-disjoint paths problem is a fundamental problem in the design and implementation of parallel and distributed computing systems and it has been extensively studied for a variety of (families of) interconnection networks. Whilst Menger's Theorem implies that, given a source node and n distinct target nodes (different from the source) in a graph of node-connectivity n, there exist n node-disjoint paths to each of the target nodes from the source, it is by no means easy to identify and actually construct the paths, especially if the paths are to be as short as possible. Indeed, given a source and a collection of target nodes in an arbitrary graph, the general problem of finding node-disjoint paths from the source to each of the target nodes with each path of shortest length is NP-hard. However, in many interconnection networks, which almost always have 'uniformity' properties such as recursive decomposability, node-symmetry, and degree regularity, the situation is much more acceptable. For example: Rabin studied the one-to-many node-disjoint paths problem in hypercubes where he showed that given a source node and n target nodes in an n-dimensional hypercube, there exist node-disjoint paths from the source to each of the target nodes such that each path has length at most 1 plus the diameter of the n-dimensional hypercube (that is, n); and Gu and Peng showed that given a source and $n - 1$ target nodes in an *n-star* graph (introduced as an improvement to hypercubes by Akers, Horel, and Krishnamurthy), there is an algorithm of time complexity $O(n^2)$ that builds $n - 1$ paths from the source to each of the target nodes such that the length of each path is at most the diameter of the n-star graph (that is, $\lfloor 3(n-1)/2 \rfloor$) plus 2. In this talk, we shall examine the one-to-many node-disjoint paths problem in (n, k)-stars, introduced by Chiang and Chen as improvements to n-star graphs.

Testing continuous distributions

Artur Czumaj
University of Warwick

We study the task of testing properties of probability distributions and our focus is on understanding the role of continuous distributions in this setting. We consider a scenario in which we have access to independent samples of a distribution D over a potentially continuous or uncountable domain. Our goal is to test whether D has a given probability distribution or it is ε-far from it (in the *statistical distance*).

We first study the question of testing if a given probability distribution is *discrete* (has finite support). The reduction to a lower bound due to Raskhodnikova et al. (FOCS'2007) shows that testing if a probability distribution D is discrete on at most N points requires $\Omega(N^{1-o(1)})$ sample complexity. We show a testing algorithm that almost matches this bound and does the test with $O(N/\varepsilon)$ samples.

It is not difficult to see that for many natural properties of continuous distributions, no testing algorithm can exist and the central objective of our study is to understand if there are any nontrivial properties that can be efficiently tested. For example, it is easy to see that there is no testing algorithm that tests if a given probability distribution on [0,1] is uniform. We show however, that if some additional information about the input distribution is known, testing uniform distribution is possible. We extend the recent result about testing uniformity for monotone distributions on Boolean n-dimensional cubes by Rubinfeld and Servedio (STOC'2005) to the case of continuous $[0,1]^n$ cubes. We show that if a probability distribution D on $[0,1]^n$ is monotone, then one can test if D is uniform with the sample complexity $O(n/\varepsilon^2)$. This result is optimal up to at most a polylogarithmic factor.

This is a joint work with M. Adamaszek (University of Warwick) and C. Sohler (University of Bonn)**.**

Fast distance multiplication of unit-Monge matrices

Alexander Tiskin
Warwick University

A matrix is called *Monge*, if its density matrix is nonnegative. Monge matrices play an important role in optimization. Distance multiplication (also known as min-plus or tropical multiplication) of two Monge matrices of size n can be performed in time $O(n^2)$. Motivated by applications to string algorithms, we introduce the following subclass of Monge matrices: A matrix is called *unit-Monge*, if its density matrix is a permutation matrix. We further restrict our attention to a natural subclass that we call simple unit-Monge matrices; under distance multiplication, such matrices form a finite aperiodic monoid with many interesting properties. Previously, we have shown that distance multiplication of simple unit-Monge matrices can be performed in time $O(n^{1.5})$. Landau conjectured in 2006 that this problem can be solved in linear time. In the current work, we give an algorithm running in time $O(n \log n)$, thus approaching Landau's conjecture within a logarithmic factor. The new algorithm implies immediate improvements in running time for a number of string and graph algorithms; in particular, we obtain an algorithm for finding a maximum clique in a circle graph in time $O(n \log^2 n)$, and a surprisingly efficient algorithm for comparing grammar-compressed strings. We conclude that unit-Monge matrices are a fascinating object and a powerful tool, that deserves further study from both the mathematical and the algorithmic viewpoints.

Biological Perspectives on Filamental Automata

Alan Gibbons* and Martyn Amos^
*** Department of Computer Science, King's College London**
^ Department of Computing and Mathematics, Manchester Metropolitan University

Filamental cellular automata (one-dimensional arrays of identical finite automata) are of significant biological relevance. They may be used to model patterns of growth and development. We explore the simplest of such systems and discover what the minimum requirements are, in terms of numbers of states and the range of communication between the component finite automata, in order to observe coordinated, self-stabilizing behaviour in the form of waves of cellular state changes along a filament. This approach is in contrast to classical work on, for example, the firing-squad problem or token passing in circularly distributed computers where the solutions are generally more complex. Rather, we seek the simplicity that may more readily occur naturally. In addition, and surprisingly, we show that even simpler component finite automata are sufficient to guarantee the self-stability of populations of filamental automata; that is, even simpler than those that guarantee self-stability in individual filamental automata. Such populations we call viable populations. We also show that these filamental automata can be generated by evolutionary algorithms, using reasonable biological assumptions.

Expending ring search algorithms and their energy-time efficiency

Ida Pu
Department of Computing, Goldsmiths, University of London

Expending ring search algorithms are used in many reactive route discovery protocols for mobile ad hoc networks and energy consumption is one of the important issues in design of such algorithms. In this talk, I will analyze the heuristics and energy-time efficiency of the algorithms under various constraints, and introduce two generic and flexible metrics for assessment of mobile ad hoc network performance in terms of energy-time efficiency. The combined effect of energy consumption and time delay is described as a trade-off in performance analysis. The application results show that the new metrics can be useful in efficiency assessment of different protocols.

New stable matching problems arising from national matching schemes in Hungary

Péter Biró
University of Glasgow

Student admissions, for both secondary schools and higher education, are organized by centralised matching schemes in Hungary. In the case of secondary schools, the program, in operation since 2000, is based precisely on the original model and algorithm of Gale and Shapley, which appears to make it unique among similar applications. The core of the algorithm is the same for the higher education scheme, established in 1985, but this model has at least three special features that are also interesting in a theoretical sense.

The first feature is the presence of ties in the system. The attempted output of the program is a so-called stable score-limit. It can be shown that the results of Gale and Shapley apply for this generalised model as well, namely, the applicant/college-oriented algorithms produce stable score-limits and these solutions are the best/worst possible stable score-limits for the applicants. We note that in this program the college-oriented algorithm was changed to the applicant-oriented version in 2007.

The second feature is the condition of lower quotas. In addition to upper quotas, here, every college may have a lower quota as well. It is possible to show that a stable solution may not exist in this case; moreover, the problem of deciding whether a stable solution exists is NP-complete in general.

The third feature is the problem of common quotas. In this case, in addition to the individual quotas of the colleges, particular sets of colleges can have common quotas. Again, we show that a stable matching may not exist under such conditions that may occur in the current model and we prove that the related decision problem is NP-complete. On the other hand we show that for nested set systems, the problem becomes solvable by a generalised version of the Gale-Shapley algorithm. Actually this structure was present in the application until 2007, when legislative changes made the problem difficult.

This is a joint work with Tamás Fleiner, Rob W. Irving and David F. Manlove.

Finding induced paths of given parity in claw-free graphs

Pim van 't Hof
PhD student
University of Durham

The *PARITY PATH* problem asks whether a given graph *G* contains both an even length and an *odd* length induced path between two specified vertices *s* and *t.* Although the *PARITY PATH* problem is NP-complete in general, it has been shown to be solvable in polynomial time for several classes of graphs. In the related *ODD PATH* and *EVEN PATH* problems, the goal is to determine whether an induced path of *odd* (respectively even) length between two specified vertices exists, and if so, to find one. We present an algorithm that solves both the *ODD PATH* and the *EVEN PATH* problem for *claw-free* graphs in polynomial time. As an immediate consequence, our algorithm solves the *PARITY PATH* problem for claw-free graphs in polynomial time.

The Expressive Power of Binary Submodular Functions

Stanislav Zivny (joint work with David Cohen and Peter Jeavons)
Oxford University

It has previously been an open problem whether all *Boolean submodular functions* can be *decomposed* into a sum of *binary submodular functions* over a possibly larger set of variables. This problem has been considered within several different contexts in computer science, including computer vision, artificial intelligence, and pseudo-Boolean optimisation. Using a connection between the expressive power of valued constraints and certain algebraic properties of functions, we answer this question negatively. Our results have several corollaries. First, we characterise precisely which submodular functions of arity 4 can be expressed by binary submodular functions. Next, we identify a novel class of submodular functions of *arbitrary* arities which can be expressed by binary submodular functions, and therefore minimised efficiently using a so-called expressibility reduction to the *Min-Cut* problem. More importantly, our results imply limitations on this kind of reduction and establish for the first time that it cannot be used in general to minimise arbitrary submodular functions. Finally, we refute a conjecture of Promislow and Young on the structure of the extreme rays of the cone of Boolean submodular functions.

Densities and Partitions

Sharon Curtis
Oxford Brookes University

In the land of bioinformatics, there lives a problem of finding a *maximally dense* segment of an input sequence. The density of a segment is the total sum of the areas of the elements in that segment divided by the total sum of their (positive) widths. This talk will look at some efficient approaches to calculating a maximum density segment in linear time.

Descriptive Complexity of Optimisation Problems

James Gate
(part of doctoral study; supervised by Professor Iain Stewart)
Department of Computer Science, University of Durham

The field of Descriptive Complexity is the bridge between finite model theory and algorithmic complexity theory. The majority of research in this field has focused on classes of decision problems and the logics that capture them. This talk looks at how to extend these logics to capture optimization problems. Specifically, it shall examine the class of (*deterministic*) polynomial time optimization problems (referred to as Popt) and argue that a single logical framework, which does not discriminate between *maximization* and *minimization* problems, is the most appropriate way to capture this class. Such a framework, using fixed-point operators along with examples of their use, shall be presented.

Computing the Factors of the l^{th} Cyclotomic Polynomial over F_p

Andrew Potter
University of Bristol

Consider the following algorithmic problem. Fix an *odd* prime *l*. Given an *odd* prime *p* != *l*, *factorise* the l^{th} *cyclotomic polynomial* over F_p, the finite field of *p* elements.

In 1990, Jonathan Pila demonstrated a deterministic polynomial-time (in log *p*) algorithm which produces a solution to this problem in the case that *p* = 1 mod *l*. Pila's work generalises an algorithm of Rene Schoof, which calculates the number of points on an elliptic curve over a finite field, and its application to finding square roots mod *p*.

It has been the intent of my research to generalise Pila's method to obtain a deterministic polynomial-time algorithm which solves the problem for all *p* != *l*. In my presentation, I will develop the necessary machinery for the method, and present the current progress of my research.

Data locality problems and colored bin packing

Andy Twigg
University of Oxford

I will talk about a problem of data locality in networks, introduced by Fan Chung et al (2006). The problem is to distribute colored items among processors so that items of each color span only few processors. I will present the first online and offline approximation schemes for the problem, and give some inapproximability results. If time permits, I will discuss a more combinatorial generalization involving hypergraph edge coverings.

Ranking and Suggesting Popular Items

James Cruise
University of Bristol

We consider the problem of ranking the popularity of items and suggesting popular items based on user feedback. User feedback is obtained by iteratively presenting a set of suggested items, and users selecting items based on their own preferences either from this suggestion set or from the set of all possible items. The goal is to quickly learn the true popularity ranking of items, and suggest popular items. The difficulty is that making suggestions to users can reinforce popularity of some items and distort the resulting ranking. We propose and study several algorithms for ranking and suggesting popular items.

Pattern matching with few errors and wildcard symbols

Raphaël Clifford
University of Bristol

We consider the classic problem of pattern matching with few mismatches in the presence of promiscuously matching wildcard symbols. Given a text t of length n and a pattern p of length m with optional wildcard symbols and a bound k, our algorithms find all the alignments for which the pattern matches the text with Hamming distance at most k and also returns the location and identity of each mismatch. The algorithms we present take $\tilde{O}(kn)$ time and are therefore optimal to within logarithmic factors if all mismatch positions are required in the output. The deterministic solutions we develop borrow from the tool set of algebraic coding theory and provide a new framework in which to tackle approximate pattern matching problems.