Report on BCTCS 2023

The 39th British Colloquium for Theoretical Computer Science

3–4 April 2023, University of Glasgow

Ciaran McCreesh

The British Colloquium for Theoretical Computer Science (BCTCS) is an annual forum in which researchers in Theoretical Computer Science can meet, present research findings, and discuss developments in the field. It also provides a welcoming environment for PhD students to gain experience in presenting their work to a broader audience, and to benefit from contact with established researchers.

BCTCS 2023 was hosted by the University of Glasgow and held from 3rd to 4th April 2023. The event attracted 37 registered participants, and featured an interesting and wide-ranging programme. A total of 15 contributed talks – predominantly by PhD students – were presented at the meeting alongside the five keynote speakers. The meeting also featured a special session on the pedagogy of theoretical computer science.

BCTCS 2024 will be hosted by the University of Bath from 3rd–5th April 2024. Researchers and PhD students wishing to contribute talks concerning any aspect of Theoretical Computer Science are cordially invited to do so. Further details are available from the BCTCS website at `www.bctcs.ac.uk`.

Invited Talks

Ruth Hoffmann (University of St Andrews)

Composable Constraint Models for Permutation Patterns and their enumeration

Permutation pattern research started off as investigating which sequences of numbers can be sorted by using a stack. This has now extended into many fields such as using permutations which contain or avoid certain types of patterns when investigating for example Mahonian statistics. Constraint programming is a way of solving combinatorial problems by taking variables, the values they can take and constraints which involve the variables. It then searches for one (all, or the optimal) solution (variable, value assignments) which does not violate the constraints. We will explore different permutation patterns, properties and statistics. While giving you the many definitions we will see how each translates into a constraint model. Having these many models means that we can now easily mix and match them into useful tools to help solve or help investigate permutation problems computationally.

Steve Linton (University of St Andrews)
Three Trips Around the "Virtuous Circle": Theory, Algorithms, Software and Experiments

My thesis in this talk is that there can be a powerful synergy between the study of mathematical and combinatorial structures in the abstract; the theoretical study of algorithms for computing with those structures and their complexity; the development of flexible and usable software implementations of those algorithms; and the gathering of experimental data using that software, which can fuel new conjectures and lead to new mathematical results, completing the circle. I will illustrate this thesis with examples from three areas: permutation groups; transformation monoids; and token-passing networks and pattern classes of permutations.

David Manlove (University of Glasgow)
Models and Algorithms for the Kidney Exchange Problem

A patient who requires a kidney transplant, and who has a willing but incompatible donor, may be able to 'swap' his or her donor with that of another patient, who is in a similar situation, in a cyclic fashion. Altruistic donors can also trigger "chains" of transplants involving multiple recipients together with their willing but incompatible donors. Kidney exchange programmes (KEPs) organise the systemic detection of optimal sets of cycles and chains based on their pools of donors and recipients. There are many examples of KEPs around the world, including the UK Living Kidney Sharing Scheme (UKLKSS). In this talk I will describe integer programming models and algorithms that can be used to solve the underlying optimisation problem involved in a KEP. This includes the algorithms developed at Glasgow that have been used by NHS Blood and Transplant for the UKLKSS since 2008.

Faron Moller (Swansea University)
Technocamps: Transforming Digital Education Throughout Wales

By 2000, it became evident that, in Wales, interest in, knowledge of, and capacity for computing was not keeping pace with the transformational rise of the digital society and economy. Technocamps, the pan-Wales school and community outreach unit established at Swansea University but with a hub in every university in Wales, has throughout this time researched, championed and delivered change in national curricula, qualifications, delivery and professional development in order to foster a sustainable digital skills pipeline in Wales. In this presentation, we highlight the activities and impact of Technocamps, showcasing its wider impact on computing education, practitioners, schools, and learners in Wales, especially with the introduction of the new Curriculum for Wales in September 2022, with its major reform of computer science and cross-curricular digital competencies.

Syed Waqar Nabi (University of Glasgow)
Navigating Pedagogies: Teaching Theory-Heavy Courses to Software Engineering Student

The landscape of teaching pedagogies is a rich one, and not always easy to navigate. While there has been a shift towards student-centered, "constructivist" approaches to teaching, the more traditional teacher-centred approaches like direct-instruction are still prevalent, more so in theory-heavy courses. For this talk, I will use theory-heavy courses I teach to Software Engineering Graduate Apprenticeship students as conduits for exploring this pedagogy landscape, discussing experiences with using a number of teaching instruments along the way. This will lead to the specific pedagogy that I have been converging on called "Competency-Based Education" (CBE), similar to what's called "mastery learning". Based on the outcome of a working group on CBE, I will go a bit more into what CBE is, how it relates to some other pedagogies, how we can draw inspiration from other teaching domains, and what it might mean to use it for computing education. Finally, I will connect this discussion to theory-heavy computing science courses in general, and share some thoughts on how (or if) CBE can work for such courses.

Contributed Talks

Andrew Ryzhikov (University of Oxford)
On Cost Register Automata with Few Registers

Cost register automata (CRA) are an extension of deterministic finite state automata. Instead of accepting or rejecting words, they assign each word a value (which can mean, for example, a cost, probability or duration of an event). This value is computed with a finite set of write-only registers which are updated every time a transition is taken. CRA are tightly related to weighted automata (WA), and natural syntactic restrictions for CRA allow to define new subclasses of functions which are not definable in terms of WA. One such restriction is to bound the number of registers. We show that for CRA with only three registers universality (are the values of all words below/above a certain threshold?) remains undecidable both over the tropical semiring and over the semiring of rational numbers with usual addition and multiplication. In contrast, we show that the zeroness problem (does there exist a word of value zero?) for CRA over the tropical semiring becomes solvable in polynomial time if the number of registers is constant, while it is PSPACE-complete without this assumption.

This is a joint work with Laure Daviaud (City, University of London).

Peter Strulo (University of Warwick)
An Exercise in Tournament Design: When Some Matches Must Be Scheduled

In single-elimination tournaments, players play one-on-one matches with the winner proceeding to the next round until only one player remains. The problem of manipulating the outcome of the tournament by carefully choosing which opponents play each other in each round (the seeding) has been studied extensively. We introduce a new variant of this problem where the aim is to choose a seeding which results in certain desired matches being played, rather than a specific player winning. We obtain both hardness and tractability results: the problem is NP-hard in general but polynomial-time solvable when the input digraph modelling the pairwise results is acyclic.

Marcel De Sena Dall'Agnol (University of Warwick)
Streaming zero-knowledge proofs

We initiate the study of zero-knowledge proofs for data streams. Streaming interactive proofs (SIPs) are well-studied protocols whereby a space-bounded algorithm with one-pass sequential access to a massive stream of data communicates with an all-powerful but untrusted prover to verify a computation that requires large space.

We define the notion of zero-knowledge in the streaming setting and construct zero-knowledge SIPs for the two main building blocks in the streaming interactive proofs literature: the sumcheck and polynomial evaluation protocols. To the best of our knowledge all known streaming interactive proofs are based on either of these tools, and indeed, this allows us to obtain zero-knowledge SIPs for central and well-studied streaming problems, such as index, frequency moments, and inner product. Our protocols are efficient both in terms of time and space, as well as communication: the space complexity is polylog(n) and, after a non-interactive setup that uses a random string of near-linear length, the remaining parameters are sub-polynomial.

En route, we develop a toolkit for designing zero knowledge data stream protocols that may be of independent interest, consisting of an algebraic streaming commitment protocol and a temporal commitment protocol. The analysis of our protocols relies on delicate algebraic and information-theoretic arguments and reductions from average-case communication complexity.

Bruno Pasqualotto Cavalar (University of Warwick)
Constant-Depth Circuits vs. Monotone Circuits

We establish strong separations between the power of monotone and general (non-monotone) Boolean circuits:

- For every $k \geq 1$, there is a monotone function in AC^0 (constant-depth poly-size circuits) that requires monotone circuits of depth $\Omega(\log^k n)$. This vastly extends a classical result of Okol'nishnikova (1982) and Ajtai and Gurevich

(1987). Our separation holds for a monotone graph property, which was unknown even in the context of AC^0 versus mAC^0.

- For every $k \geq 1$, there is a monotone function in $\mathsf{AC}^0[\oplus]$ (constant-depth poly-size circuits extended with parity gates) that requires monotone circuits of size $\exp(\Omega(\log^k n))$. This makes progress towards a question posed by Grigni and Sipser (1992).

These results show that constant-depth circuits can be considerably more efficient than monotone circuits when computing monotone functions.

In the opposite direction, we observe that non-trivial simulations are possible in the absence of parity gates: every monotone function computed by an AC^0 circuit of size s and depth d can be computed by a monotone circuit of size $2^{n-n/O(\log s)^{d-1}}$. We show that the existence of significantly stronger monotone simulations would lead to breakthrough circuit lower bounds. In particular, if every monotone function in AC^0 admits a polynomial size monotone circuit, then NC^2 is not contained in NC^1.

Finally, we revisit our separation result against monotone circuit size and investigate the limits of our approach, which is based on a monotone lower bound for constraint satisfaction problems established by Göös et al. (2019) via lifting techniques. Adapting results of Schaefer (1978) and Allender et al. (2009), we obtain a classification of the monotone circuit complexity of Boolean-valued CSPs via their polymorphisms. This result and the consequences we derive from it might be of independent interest.

Nathan Flaherty (University of Liverpool)
On Transposition Distance of Words with Fixed Parikh Vectors

The operation of transposition is a permutation that swaps any two symbols in a word. The Parikh Vector P denotes the number of occurrences of each letter in a given word and the operation of transposition preserves its Parikh Vector. We consider the configuration graph where the set of vertices are words with the same Parikh Vector and edges are defined by transposition operations on these words. The question about the maximal shortest path between two words by transposition corresponds to the estimation of the diameter D in a configuration graph. We show the tight bound on the diameter D which is equal to $n - \max_{i \in [q]} P_i$ where q is the size of a finite alphabet, and $n = \sum_{i \in [q]} P_i$ is the length of considered words. The lower bound is based on the analysis of cyclic covers of auxiliary graph structure and the matching upper bound follows from the direct proof of algorithmic transformation. This is the joint work with Duncan Adamson, Igor Potapov and Paul Spirakis.

Ben Lloyd-Roberts (Swansea University)

Mining Invariants from State Space Observations

The application of model checking to verify railway signalling systems has a long history within academia and is beginning to seen some real applications in the railway sector. One limitation of such model checking is that verification can fail due to over approximation, typically when using techniques such as inductive verification. Here, one solution is to introduce so-called invariants, formal properties satisfied by all states, to suppress false positives. However, automatically generating sufficiently strong invariants to help bound the region of reachable states is complex. In this work, we show it is possible to use machine learning to generate candidate invariants for model checking. Our methodology starts by providing a first formal mapping of state spaces to a reinforcement learning (RL) environment. We then train agents to explore large regions of states spaces while building a dataset of unique state observations. Finally we demonstrate that statistical analysis of state observations gives rise to interesting correlations between variables, allowing proposals for candidate invariant properties.

Matthew McIlree (University of Glasgow)
How can a constraint solver prove it is telling the truth?

A proof log for a problem-solving algorithm provides a verifiable certificate that the result is correct, and also an auditable record of the steps taken to obtain that result. In the field of Boolean satisfiability, proof-logging has become an expected capability of modern solvers, with a standard proof format called DRAT widely accepted for independent verification. In contrast, a similar standard practice has not yet been adopted for Constraint Programming (CP), due to the difficulties of applying DRAT to the more expressive formulations and reasoning present in this paradigm. However, recent work towards "An Auditable Constraint Programming Solver" (Gocht et al. 2022) has shown how a proof system working in a pseudo-Boolean format can certify the reasoning carried out for several important expressive global constraints, offering a strong candidate for a complete, general CP proof logging method. This talk will be an introduction to proof logging in the context of constraint programming. It will summarise the main motivations; the core techniques developed so far; and the reasons for being optimistic about the applicability of the method going forward.

Laura Larios-Jones (University of Glasgow)
Minimising temporal reachability in graphs with uncertainty

Temporal graphs consist of an underlying graph and an assignment of timesteps to edges that specifies when each edge is active. This allows us to model spread through a network which is time-sensitive. Previous work has explored minimising spread by edge deletion for applications such as epidemiology. In reality, these models cannot be exact. This motivates the introduction of uncertainty to the problem. Our goal is to remove a set of edges in our graph such that the maximum number of vertices reachable from any starting vertex is minimised even when there is uncertainty in our input. We will discuss some preliminary analogous structural and algorithmic results.

Fabricio Mendoza Granada (University of Glasgow)
On finding the b-chromatic number of a tree

Graph colouring is an extensively studied problem in computer science, discrete mathematics and other disciplines. It involves assigning colours or labels to the vertices so that not two adjacent vertices share the same colour. This assignment is called a proper colouring. The problem was originally proposed as a puzzle to colour the map of counties in England in 1878. Its applications arise in the context of scheduling, timetable construction, register allocation and many others. The first graph colouring parameter to be studied was the chromatic number of a graph G, $\chi(G)$, which is the minimum number of colours used by a proper colouring of G. In this talk we will discuss the b-chromatic number $\varphi(G)$ of a graph G, a concept introduced by Irving and Manlove in 1998. The b-chromatic number of a graph is the maximum integer k for which G admits a proper colouring such that for every colour c there exists a vertex v of colour c that is adjacent to at least one vertex of every other colour. Deciding whether $\varphi(G) \geq k$ for a given graph G and integer k was proved to be NP-complete by Irving and Manlove, and this holds even for bipartite graphs. However, they proved that the b-chromatic number of a tree can be computed in polynomial time by describing an algorithm to find a b-chromatic colouring using $\varphi(G)$ colours. In this talk we will present the algorithm for finding a b-chromatic number of a tree in pseudocode form. Futhermore, we show that the algorithm runs in linear time; previously it was only claimed that the algorithm runs in polynomial time. Finally, we will present some experimental results on families of random trees.

Filippos Pantekis (Swansea University)
GPGPUs, Supercomputers, and a Game of Chess

The evolution of General Purpose Graphics Processing Unit (GPGPU) devices, paired with their wide commercial availability, has enabled a broader spectrum of problems to benefit from the superior mathematical capabilities offered by this hardware. Perhaps the biggest obstacle in using GPGPUs to accelerate the solving for all problems, is the restrictive computation flow (regularity) expected by the hardware in order to maximise performance. This talk presents how certain algorithmic choices together with hardware-specific optimisations can transform an irregular algorithm for the N-Queens problem to a competitive solver.

Carlos A. Perez Delgado (University of Kent)
Towards a Physical Fundamental Computational Complexity Theory

In theoretical computer science, the fundamental yardstick of computational cost is left largely undefined. The (time) cost of performing a computation/algorithm is measured in its number of "primitive operations". However, one is allowed to choose the set of primitive operations at will. Big "Oh" notation allows one to do so, while retaining a consistent measure. This allows for an elegantly simple theory that can nevertheless make meaningful statements about algorithms.

The theory is not without its flaws, however. The first is the already mentioned arbitrariness of the yardstick(s). Second, it fails to say anything meaningful when attempting to compare different architectures, or comparing algorithms across them. For instance, computations running on massively parallel architectures, quantum computers, and/or single-core processors cannot be meaningfully compared with one another without introducing extra assumptions.

In this talk I will propose a fundamental theory of computational complexity. This theory uses the physical resource action (that is energy in joules times time in seconds), as the fundamental unit of computation. We will introduce a model of computation that allows us apply this metric, much like Turing machines can be used for (traditional) computational complexity cost. We will discuss how to recover all existing results from computational complexity, and we will discuss benefits of this model in terms of meaningful comparisons that traditional complexity theory cannot make.

Peace Ayegba (University of Glasgow)
Resolving the complexity of variants of stable matching problems.

Matching problems involve the allocation of one set of agents to another set of agents based on preferences, with application in various real-world centralised matching schemes. A common objective is to find a stable matching where no set of agents would rather be matched together than with their current assignment. It is well known that finding a maximum cardinality stable matching for several matching problems such as the Stable Marriage problem with Ties and Incomplete lists (MAX-SMTI), is NP-hard, even with strong restrictions on the input. However, a polynomial-time algorithm exists for a restricted version of MAX-SMTI, where each man's list is of length at most 2 and each woman's list can be of unbounded length. This talk resolves the complexity of other maximum cardinality stable matching problems (e.g., in the context of hospitals-residents with ties, and student-project allocation) under strong restrictions on the input.

Xin Ye (Durham University)
Computing Balanced Solutions for Large International Kidney Exchange Schemes

To overcome incompatibility issues, kidney patients may swap their donors. In international kidney exchange programmes (IKEPs), countries merge their national patient-donor pools. We consider a recently introduced credit system. In each round, countries are given an initial "fair" allocation of the total number of kidney transplants. This allocation is adjusted by a credit function yielding a target allocation. The goal is to find a solution that approaches the target allocation as closely as possible, to ensure long-term stability of the international pool. As solutions, we use maximum matchings that lexicographically minimize the country deviations from the target allocation. We first introduce a novel approach for incorporating credits that has not been proposed in the literature before. Namely, let the solution concepts prescribe a set of target allocations for a credit-adjusted game, where the credits are incorporated into the value function of the game directly. We perform a computational study for a large number of countries, up to fifteen countries. For the initial allocations we extend by also considering the tau value and Banzhaf value, and compare them to previously obtained results, namely the benefit value, contribution value, Shapley value and nucleolus. Our experiments show that using lexicographically minimal maximum matchings instead of ones that only minimize the largest deviation from the target allocation (as previously done) may make an IKEP up to 54% more balanced. This is joint work with Marton Benedek, Peter Biro and Daniel Paulusma.

David Kutner (Durham University)
The TaRDiS and epidemics in temporal graphs

We are interested in the resilience to infection of a population of n individuals who will interact m times, with k of those individuals being initially infectious. In the worst case, the entire population is infected once all the interactions have occured; and this is necessarily the case $k = n$. Our question is then to find the size of the smallest set of infectious individuals which would still infect the entire population.

Temporal graphs (graphs which change over time) offer us a convenient model for this problem. In our case, vertices corresponding to individuals remain constant and edges, each corresponding to an interaction, appear at exactly one (unique) time. We denote this temporal graph $G = (V, E), \lambda : E \to [|E|]$, and say a node v_0 reaches another node v_l (denoted by $v_0 \rightsquigarrow v_l$) if there is a static path $v_0, v_1, \ldots, v_l \in G$ and $\lambda(v_i, v_{i+1})\ \forall i \in \{0, l-1\}$.

Then our problem can be formalized as follows: given a temporal graph G, λ, and integer k, is there a set of vertices $S \subseteq V(G)$ such that for every $v \in V(G)$ either $v \in S$ or $\exists u \in S : u \rightsquigarrow v$? We call this problem Temporal Reachability Dominating Set, or TaRDiS, and present hardness and tractability results for it.

Thomas Karam (University of Oxford)

Lower-order ranks and the structure of the ranges of boolean polynomials on finite prime fields

Let p be a prime integer, and let $1 \leq d < p$ be a positive integer. The *degree-d rank* of a polynomial $P : \mathbb{F}_p^n \to \mathbb{F}_p$ was defined in 2007 by Green and Tao as the smallest nonnegative integer k such that we can find polynomials $P_1, \dots, P_k : \mathbb{F}_p^n \to \mathbb{F}_p$ with degree at most d and a function $F : \mathbb{F}_p^k \to \mathbb{F}_p$ satisfying $P = F(P_1, \dots, P_k)$.

As shown by Green and Tao, if $2 \leq d < p$ and P is a degree-d polynomial not approximately uniformly distributed on $\mathbb{F}_p^n$, then P must have bounded degree-$(d-1)$ rank. The literature after that has largely focused on the equidistribution of polynomials, and hence on the notion of degree-$(d-1)$ rank.

Recently, Gowers and the speaker showed that this statement could be extended to boolean polynomials: if P is not approximately uniformly distributed on $\{0,1\}^n$, then P coincides on $\{0,1\}^n$ with a polynomial that has bounded degree-$(d-1)$ rank.

In this talk I shall explain how this extension, which is still based purely on the degree-$(d-1)$ rank, may be used to deduce a description of the range of a polynomial on S^n which uses the other ranks defined by Green and Tao.