Friday, June 21

	Fri-10A: Data Structures II	Fri-10B: Graph Drawing II
9:00-9:20	**A New Lower Bound for Semigroup Orthogonal Range Searching** Peyman Afshani We report the first improvement in the space-time trade-off of lower bounds for the orthogonal range searching problem in the semigroup model, since Chazelle's result from 1990. This is one of the very fundamental problems in range searching with a long history. Previously, Andrew Yao's influential result had shown that the problem is already non-trivial in one dimension [*Space-time tradeoff for answering range queries*, STOC 1982]: using m units of space, the query time $Q(n)$ must be $\Omega(\alpha(m,n)+\frac{n}{m-n+1})$ where $\alpha(\cdot,\cdot)$ is the inverse Ackermann's function, a very slowly growing function. In d dimensions, Bernard Chazelle [*Lower bounds for orthogonal range searching: part II. the arithmetic model*, JACM 1990] proved that the query time must be $Q(n)=\Omega((\log_\beta n)^{d-1})$ where $\beta = 2m/n$. Chazelle's lower bound is known to be tight for when space consumption is "high" i.e., $m=\Omega(n\log^{d+\varepsilon} n)$. We have two main results. The first is a lower bound that shows Chazelle's lower bound was not tight for "low space": we prove that we must have $mQ(n)=\Omega(n(\log n\log\log n)^{d-1})$. Our lower bound does not close the gap to the existing data structures, however, our second result is that our analysis is tight. Thus, we believe the gap is in fact natural since lower bounds are proven for idempotent semigroups while the data structures are built for general semigroups and thus they cannot assume (and use) the properties of an idempotent semigroup. As a result, we believe to close the gap one must study lower bounds for non-idempotent semigroups or building data structures for idempotent semigroups. We develope significantly new ideas for both of our results that could be useful in pursuing either of these directions.	**Dual Circumference and Collinear Sets** V. Dujmović and P. Morin We show that, if an n-vertex triangulation T of maximum degree Δ has a dual that contains a cycle of length ℓ, then T has a non-crossing straight-line drawing in which some set, called a *collinear set*, of $\Omega(\ell/\Delta^4)$ vertices lie on a line. Using the current lower bounds on the length of longest cycles in 3-regular 3-connected graphs, this implies that every n-vertex planar graph of maximum degree Δ has a collinear set of size $\Omega(n^{0.8}/\Delta^4)$. Very recently, Dujmović *et al* (SODA 2019) showed that, if S is a collinear set in a triangulation T then, for any point set $X\subset\mathbb{R}^2$ with $\lvert X\rvert = \lvert S\rvert$, T has a non-crossing straight-line drawing in which the vertices of S are drawn on the points in X. Because of this, collinear sets have numerous applications in graph drawing and related areas.
9:20-9:40	**Independent Range Sampling, Revisited Again** Peyman Afshani and Jeff M. Phillips We revisit the range sampling problem: the input is a set of points where each point is associated with a real-valued weight. The goal is to store them in a structure such that given a query range and an integer k, we can extract k independent random samples from the points inside the query range, where the probability of sampling a point is proportional to its weight. This line of work was initiated in 2014 by Hu, Qiao, and Tao and it was later followed up by Afshani and Wei. The first line of work mostly studied unweighted but dynamic version of the problem in one dimension whereas the second result considered the static weighted problem in one dimension as well as the unweighted problem in 3D for halfspace queries. We offer three main results and some interesting insights that were missed by the previous work: We show that it is possible to build efficient data structures for range sampling queries if we allow the query time to hold in expectation (the first result), or obtain efficient worst-case query bounds by allowing the sampling probability to be approximately proportional to the weight (the second result). The third result is a conditional lower bound that shows essentially one of the previous two concessions is needed. For instance, for the 3D range sampling queries, the first two results give efficient data structures with near-linear space and polylogarithmic query time whereas the lower bound shows with near-linear space the worst-case query time must be close to $n^{2/3}$, ignoring polylogarithmic factors. Up to our knowledge, this is the first such major gap between the expected and worst-case query time of a range searching problem.	**Cubic Planar Graphs That Cannot Be Drawn On Few Lines** David Eppstein For every integer ℓ, we construct a cubic 3-vertex-connected planar bipartite graph G with $O(\ell^3)$ vertices such that there is no planar straight-line drawing of G whose vertices all lie on ℓ lines. This strengthens previous results on graphs that cannot be drawn on few lines, which constructed significantly larger maximal planar graphs. We also find apex-trees and cubic bipartite series-parallel graphs that cannot be drawn on a bounded number of lines.

9:40-10:00	**Dynamic Geometric Data Structures via Shallow Cuttings** T. M. Chan We present new results on a number of fundamental problems about dynamic geometric data structures: 1. We describe the first fully dynamic data structures with sublinear amortized update time for maintaining (i) the number of vertices or the volume of the convex hull of a 3D point set, (ii) the largest empty circle for a 2D point set, (iii) the Hausdorff distance between two 2D point sets, (iv) the discrete 1-center of a 2D point set, (v) the number of maximal (i.e., skyline) points in a 3D point set. The update times are near $n^{11/12}$ for (i) and (ii), $n^{7/8}$ for (iii) and (iv), and $n^{2/3}$ for (v). Previously, sublinear bounds were known only for restricted "semi-online" settings [Chan, SODA 2002]. 2. We slightly improve previous fully dynamic data structures for answering extreme point queries for the convex hull of a 3D point set and nearest neighbor search for a 2D point set. The query time is $O(\log^2 n)$, and the amortized update time is $O(\log^4 n)$ instead of $O(\log^5 n)$ [Chan, SODA 2006; Kaplan et al., SODA 2017]. 3. We also improve previous fully dynamic data structures for maintaining the bichromatic closest pair between two 2D point sets and the diameter of a 2D point set. The amortized update time is $O(\log^4 n)$ instead of $O(\log^7 n)$ [Eppstein 1995; Chan, SODA 2006; Kaplan et al., SODA 2017].	**Connecting the Dots (with Minimum Crossings)** Akanksha Agrawal, Grzegorz Guśpiel, Jayakrishnan Madathil, Saket Saurabh, Meirav Zehavi We study a prototype Crossing Minimization problem, defined as follows. Let $\mathcal{F}$ be an infinite family of (possibly vertex-labeled) graphs. Then, given a set P of (possibly labeled) n points in the Euclidean plane, a collection $L \subseteq \mathsf{Lines}(P) = \{\ell : \ell$ is a line segment with both endpoints in $P\}$, and a non-negative integer k, decide if there is a sub-collection $L' \subseteq L$ such that the graph $G = (P, L')$ is isomorphic to a graph in $\mathcal{F}$ and L' has at most k crossings. By $G = (P, L')$, we refer to the graph on vertex set P, where two vertices are adjacent if and only if there is a line segment that connects them in L'. Intuitively, in Crossing Minimization, we have a set of locations of interest, and we want to build/draw/exhibit connections between them (where L indicates where it is feasible to have these connections) so that we obtain a structure in $\mathcal{F}$. Natural choices for $\mathcal{F}$ are the collections of perfect matchings, Hamiltonian paths, and graphs that contain an (s, t)-path (a path whose endpoints are labeled). While the objective of seeking a solution with few crossings is of interest from a theoretical point of view, it is also well motivated by a wide range of practical considerations. For example, links/roads (such as highways) may be cheaper to build and faster to traverse, and signals/moving objects would collide/interrupt each other less often. Further, graphs with fewer crossings are preferred for graphic user interfaces. As a starting point for a systematic study, we consider a special case of Crossing Minimization. Already for this case, we obtain NP-hardness and W[1]-hardness results, and ETH-based lower bounds. Specifically, suppose that the input also contains a collection D of d non-crossing line segments such that each point in P belongs to exactly one line in D, and L does not contain line segments between points on the same line in D. Clearly, Crossing Minimization is the case where $d = n$—then, P is in general position. The case of $d = 2$ is of interest not only because it is the most restricted non-trivial case, but also since it corresponds to a class of graphs that has been well studied—specifically, it is Crossing Minimization where $G = (P, L)$ is a (bipartite) graph with a so called *two-layer drawing*. For $d = 2$, we consider three basic choices of $\mathcal{F}$. For perfect matchings, we show *(i)* NP-hardness with an ETH-based lower bound, *(ii)* solvability in subexponential parameterized time, and *(iii)* existence of an $O(k^2)$-vertex kernel. Second, for Hamiltonian paths, we show *(i)* solvability in subexponential parameterized time, and *(ii)* existence of an $O(k^2)$-vertex kernel. Lastly, for graphs that contain an (s, t)-path, we show *(i)* NP-hardness and W[1]-hardness, and *(ii)* membership in XP.
10:00-10:30	Coffee Break	
	Fri-11A: Complexity	**Fri-11B: Combinatorial Geometry III**
10:30-10:50	**The Unbearable Hardness of Unknotting** A. de Mesmay, Y. Rieck, E. Sedgwick, M. Tancer We prove that deciding if a diagram of the unknot can be untangled using at most k Reidemeister moves (where k is part of the input) is **NP**-hard. We also prove that several natural questions regarding links in the 3-sphere are **NP**-hard, including detecting whether a link contains a trivial sublink with n components, computing the unlinking number of a link, and computing a variety of link invariants related to four-dimensional topology (such as the 4-ball Euler characteristic, the slicing number, and the 4-dimensional clasp number).	**An Experimental Study of Forbidden Patterns in Geometric Permutations by Combinatorial Lifting** Goaoc X., Holmsen A., and Nicaud C. We study the problem of deciding if a given triple of permutations can be realized as geometric permutations of disjoint convex sets in $\mathbb{R}^3$. We show that this question, which is equivalent to deciding the emptiness of certain semi-algebraic sets bounded by cubic polynomials, can be "lifted" to a purely combinatorial problem. We propose an effective algorithm for that problem, and use it to gain new insights into the structure of geometric permutations.

10:50-11:10	**Circumscribing Polygons and Polygonizations for Disjoint Line Segments** H. A. Akitaya, M. Korman, M. Rudoy, C. D. Tóth, and D. L. Souvaine Given a planar straight-line graph $G = (V, E)$ in $\mathbb{R}^2$, a *circumscribing polygon* of G is a simple polygon P whose vertex set is V, and every edge in E is either an edge or an internal diagonal of P. A circumscribing polygon is a *polygonization* for G if every edge in E is an edge of P. We prove that every arrangement of n disjoint line segments in the plane has a subset of size $\Omega(\sqrt{n})$ that admits a circumscribing polygon, which is the first improvement on this bound in 20 years. We explore relations between circumscribing polygons and other problems in combinatorial geometry, and generalizations to $\mathbb{R}^3$. We show that it is NP-complete to decide whether a given graph G admits a circumscribing polygon, even if G is 2-regular. Settling a 30-year old conjecture by Rappaport, we also show that it is NP-complete to determine whether a geometric matching admits a polygonization.	**A Product Inequality for Extreme Distances** Adrian Dumitrescu Let $p_1, \ldots, p_n$ be n distinct points in the plane, and assume that the minimum inter-point distance occurs $s_{\min}$ times, while the maximum inter-point distance occurs $s_{\max}$ times. It is shown that $s_{\min} s_{\max} \leq \frac{9}{8} n^2 + O(n)$; this settles a conjecture of Erdős and Pach (1990).
11:10-11:30	**Counting Polygon Triangulations is Hard** David Eppstein We prove that it is #P-complete to count the triangulations of a (non-simple) polygon.	**Convex Polygons in Cartesian Products** J.-L. De Carufel, A. Dumitrescu, W. Meulemans, T. Ophelders, C. Pennarun, C. D. Tóth, and S. Verdonschot We study several problems concerning convex polygons whose vertices lie in a Cartesian product of two sets of n real numbers (for short, *grid*). First, we prove that every such grid contains a convex polygon with $\Omega(\log n)$ vertices and that this bound is tight up to a constant factor. We generalize this result to d dimensions (for a fixed $d \in \mathbb{N}$), and obtain a tight lower bound of $\Omega(\log^{d-1} n)$ for the maximum number of points in convex position in a d-dimensional grid. Second, we present polynomial-time algorithms for computing the longest convex polygonal chain in a grid that contains no two points with the same x- or y-coordinate. We show that the maximum size of such a convex polygon can be efficiently approximated up to a factor of 2. Finally, we present exponential bounds on the maximum number of convex polygons in these grids, and for some restricted variants. These bounds are tight up to polynomial factors.
11:30-11:40	Break	

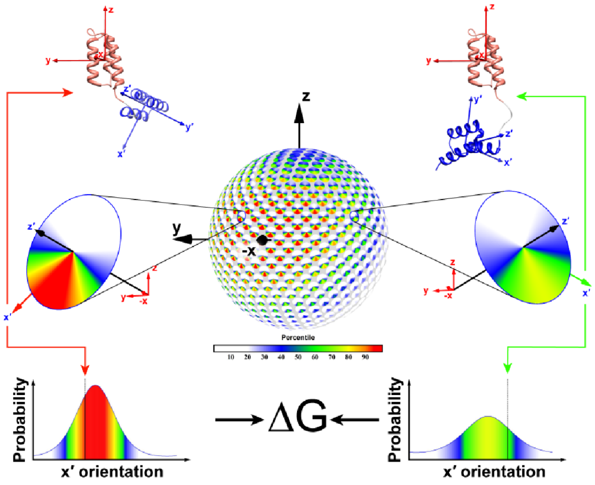
See next page ...

11:40-12:40

Invited Talk: **Some geometric and computational challenges arising in structural molecular biology**
Bruce R. Donald

Abstract: Computational protein design is a transformative field with exciting prospects for advancing both basic science and translational medical research. New algorithms blend discrete and continuous geometry to address the challenges of creating designer proteins. I will discuss recent progress in this area and some interesting open problems.

I will motivate this talk by discussing how, by using continuous geometric representations within a discrete optimization framework, broadly-neutralizing anti-HIV-1 antibodies were computationally designed that are now being tested in humans – the designed antibodies are currently in eight clinical trials (See `https://clinicaltrials.gov/ct2/results?cond=&term=vrc07&cntry=&state=&city=&dist=`), one of which is Phase 2a (NCT03721510). These continuous representations model the flexibility and dynamics of biological macromolecules, which are an important structural determinant of function.



However, reconstruction of biomolecular dynamics from experimental observables requires the determination of a conformational probability distribution. These distributions are not fully constrained by the limited geometric information from experiments, making the problem ill-posed in the sense of Hadamard. The ill-posed nature of the problem comes from the fact that it has no unique solution. Multiple or even an infinite number of solutions may exist. To avoid the ill-posed nature, the problem must be regularized by making (hopefully reasonable) assumptions.

I will present new ways to both represent and visualize correlated inter-domain protein motions (See Figure). We use Bingham distributions, based on a quaternion fit to circular moments of a physics-based quadratic form. To find the optimal solution for the distribution, we designed an efficient, provable branch-and-bound algorithm that exploits the structure of analytical solutions to the trigonometric moment problem. Hence, continuous conformational PDFs can be determined directly from NMR measurements. The representation works especially well for multi-domain systems with broad conformational distributions.

Ultimately, this method has parallels to other branches of geometric computing that balance discrete and continuous representations, including physical geometric algorithms, robotics, computational geometry, and robust optimization. I will advocate for using continuous distributions for protein modeling, and describe future work and open problems.

Department of Computer Science, Department of Chemistry, and Department of Biochemistry Duke University, and Duke University Medical Center (`www.cs.duke.edu/brd/`)