CCC
Computing Community Consortium
Catalyst

# Theoretical Computer Science: Foundations for an Algorithmic World

*A Computing Community Consortium (CCC) Quadrennial Paper*

*Shuchi Chawla (University of Wisconsin-Madison), Jelani Nelson (University of California, Berkeley), Chris Umans (California Institute of Technology), and David Woodruff (Carnegie Mellon University)*

## Introduction

Theoretical Computer Science (TCS) is the subdiscipline of Computer Science that studies computational and algorithmic processes and interactions. Work in this field is mathematical rather than empirical, and offers a unique perspective that excels at exposing and answering key questions about the possibilities and limitations of computation, broadly construed. Whereas differential equations model the physical world, *algorithms* and *computation* are the language for describing, understanding, and manipulating our computational world. Theoretical Computer Science forms the scientific foundation for the study of algorithms and computation.

Theoretical Computer Science impacts computing and society by identifying key issues in new areas and framing them in ways that drive development. In fact much of the history of Computer Science, as viewed through the lens of Turing Award citations, is filled with examples of major fields that were pioneered by TCS researchers: cryptography (Adleman, Rivest, Shamir, Micali, Goldwasser); the modern theory of algorithms and computational complexity (Cook, Karp, Hopcroft, Tarjan, Hartmanis, Stearns, Blum, Yao); the foundations of machine learning (Valiant); and distributed systems (Lamport, Liskov). More recently, TCS has played a central role in the creation of the fields of quantum computation, algorithmic economics, algorithmic privacy and algorithmic fairness.

A unique feature of Theoretical Computer Science is its ability to discern computation and algorithms in settings beyond Computer Science proper. Examples include neuroscience, where one models systems within the brain and nervous system as computation by networks of discrete elements (for a seminal example of this perspective, see Valiant's "Circuits of the Mind"); systems biology, where networks of interacting genes or proteins are understood via computational models; economics, where self-interested interacting entities can be naturally seen to be performing a distributed computation; and physics, where the fundamental notions of information and computation, and insights from TCS, are deeply entwined in the current frontiers of research in gravity and quantum mechanics. Breakthrough results in pure mathematics and statistics by TCS researchers are becoming increasingly common (for example, the refutation of the Connes Embedding Conjecture, and the proof of the Kadison-Singer Conjecture).

Key technologies that grew out of Theoretical Computer Science have had a major impact in industry. Examples include the discovery of the principles that led to Google's PageRank algorithm; the development of Consistent Hashing, which in large part spawned Akamai; various key innovations in coding theory such as rateless expander codes for fast streaming, polar codes as part of the 5G standard, and local reconstruction codes for cloud storage; fast, dynamic algorithms that form the basis of navigation systems such as Google Maps and Waze, and other applications; quantum computing; cryptographic innovations such as public-key encryption and multiparty computation that form the foundation of a secure Internet; and the cryptographic underpinnings of cryptocurrencies and blockchain technology.

Despite this wide range of application areas and intersecting scientific disciplines, the field of TCS has thrived on the basis of a strong identity, a supportive community, and a unifying scientific aesthetic. The field is defined by a well-developed common language and approach that exposes previously unknown connections between disparate topics, thereby driving progress.

This document presents the case for robust support of foundational work in TCS, structured so as to allow unfettered exploration guided by the opportunities and needs of a rapidly changing and dynamic field, and thereby bringing about the greatest possible impact to society. This model has been extraordinarily successful in the past. We also highlight three major areas of current interest as possible targets for more focused support, under the TCS umbrella.

## A Strong Theoretical Computer Science Foundation

The core of TCS encompasses some of the deepest and most fascinating questions in Computer Science and Mathematics, including the P vs. NP problem and many other fundamental questions about the possibilities and limitations of algorithms and computation. But it is also teeming with more modern problems that arise from an expansive view of computation as a language for understanding systems in many contexts. TCS has a unique way of framing these problems in terms of tradeoffs between computational resources; for example, between optimality and computational efficiency, online versus offline computation, centralized versus distributed equilibria, or degree of interaction versus security. Important fields such as streaming algorithms, rigorous cryptography, and algorithmic privacy were the direct outgrowth of this “TCS approach”. Core TCS research focuses on a computational **problem** as the central object, rather than the development of prevailing algorithms and methods for a given problem. This viewpoint drives innovation and has led to unconventional approaches and solutions, often by identifying previously unknown connections between fields.

Importantly, these developments and many others arise from a **strongly supported core of TCS researchers** adept at identifying important problems from a broad range of settings, applying the tools of TCS, pulling in (and in some cases developing) sophisticated mathematics, and following the research wherever it leads.

Sustained and predictable investment in core TCS, supporting the best ideas wherever they arise, is key to continued innovation and success in the coming decade. While it is difficult to predict which advances will have the widest impact, past experience shows that investing in core TCS produces profound returns.

## Some Highlighted Areas for Focused Support Within TCS

TCS is poised to make significant contributions in several impactful directions in the coming decade including, for example, data-driven computation, algorithmic foundations of the modern economy, secure computation, applications of quantum computing, and the theory of deep learning. In this document we highlight three specific areas of impact. A forthcoming, longer "Visions for TCS" document provides an in-depth discussion of other areas of potential impact.

1. **Algorithmic Privacy:** Over the past decade our ever-increasing ability to process large amounts of data has greatly expanded our computational capabilities. However it has also brought into focus the challenges in using data without revealing sensitive information about the individuals to whom the data pertains to. Rapidly growing privacy concerns have the potential to derail progress in crucial application areas such as healthcare. The notion of Differential Privacy arose from work in TCS as a rigorous mathematical framework to quantify privacy loss and to develop techniques balancing privacy requirements with preserving the utility of the datasets. This framework has been adopted by many players in industry and government, including Apple, Google, and the US Census Bureau. However, many challenges remain unsolved including, for example, safe sharing of sensitive data across different agencies. Differential Privacy and private data analysis remain rapidly-growing fields with opportunities for both theoretical development and diverse applications.
2. **Algorithmic Fairness:** Algorithms are increasingly informing decisions deeply intertwined in our lives, from news article recommendations to criminal sentencing decisions and healthcare diagnostics. This progress, however, raises (and is impeded by) a host of concerns regarding the societal impact of computation. A prominent concern is that these algorithms, and the data underlying them, exacerbate social biases and unfairness. A nascent line of work within TCS investigates from a computational perspective how fairness should be defined and enforced, and aims to bring mathematical rigor and provable guarantees to this area. It also aims to develop techniques for efficiently auditing and certifying the trustworthiness of society-facing algorithms. Unfettered growth in the use of computing without the guardrails provided by this framework will threaten public trust in computational systems. The eventual goal of this line of work is to bring about a new algorithmic revolution where the full promise of computational technology can be realized alongside guarantees for fairness.
3. **Foundations of Data Science:** Data Science is a discipline focused on extracting knowledge from data using tools from mathematics, optimization, statistics, and computer science. Though the term was introduced more than 40 years ago, it only became widely used in the early to mid 2000s, coinciding with tremendous advances in our ability to collect and process large amounts of data. TCS has been at the forefront of this data revolution of the past decade, with major contributions to the foundations of adaptive data analysis, minimizing bandwidth in distributed data processing and developing succinct synopses of data ("sketching"), efficient algorithms for statistical problems such as robust parameter estimation, sublinear time algorithms, and rigorous understanding of problems in Machine Learning and AI. Although this data revolution has exhibited spectacular results in many applications, its continued success faces many challenges from, for example, noisy data, malicious

data, application domains with ever-increasing complexity, and more. TCS can provide a fundamental understanding of why and how data-driven technologies work, thereby addressing these challenges head-on. The TCS perspective can be expected to continue to drive innovation in Data Science, and is important to achieving its full potential.

In conclusion, we believe that maintaining strength in core TCS through continued support and funding is vital for achieving the full promise of computing technology.

*This white paper is part of a series of papers compiled every four years by the CCC Council and members of the computing research community to inform policymakers, community members and the public on important research opportunities in areas of national priority. The topics chosen represent areas of pressing national need spanning various subdisciplines of the computing research field. The white papers attempt to portray a comprehensive picture of the computing research field detailing potential research directions, challenges and recommendations.*

*This material is based upon work supported by the National Science Foundation under Grant No. 1734706. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.*

*For citation use: Chawla S., Nelson J., Umans C. & Woodruff D. (2020) Theoretical Computer Science: Foundations for an Algorithmic World.*
*https://cra.org/ccc/resources/ccc-led-whitepapers/#2020-quadrennial-papers*