In Memoriam

Suhas Diggavi, Fellow, IEEE | University of California Los Angeles, Los Angeles, CA 90095 USA | Email: suhas@ee.ucla.edu
Abbas El Gamal, Life Fellow, IEEE | Stanford University, Stanford, CA 94305 USA | Email: abbas@ee.stanford.edu
Ioannis Kontoyiannis, Fellow, IEEE | Cambridge University, CB2 1TN Cambridge, U.K. | Email: ik355@cam.ac.uk
Ajit Paranjpe | Veeco Instruments Inc., Joy Thomas Foundation (JTF), Somerset, NJ 08873 USA | Email: ajit@joythomasfoundation.org
Murali Subbarao | ServiceNow, Joy Thomas Foundation (JTF), Saratoga, CA 95070 USA | Email: murali@joythomasfoundation.org

Joy Thomas: Legacy, Foundation and the IT Society

Abstract—In this article, we commemorate Joy Thomas, who sadly passed away in 2020. Joy is best known for his seminal textbook *Elements of Information Theory*, coauthored with his Ph.D. advisor, the late Thomas Cover. We also catalog some of Joy's technical contributions to information theory and to its practice. Joy also made an impact on industry through several successful startups. In recognition of Joy's contribution to the exposition of information theory, the IEEE Information Theory Society Board of Governors has agreed to establish the Joy Thomas IEEE Information Theory Society Tutorial Paper Award. The Joy Thomas Foundation, established to honor Joy's legacy of excellence and giving back, will sponsor this award.

Joy Thomas passed away on September 28, 2020, at the young age of 57. Joy is best known for his seminal textbook *Elements of Information Theory*, coauthored (as a graduate student) with his Ph.D. advisor the late Thomas Cover [1], [2]. His book literally rewrote the way information theory is taught, making it much more accessible not only to graduate students interested in communication and compression, but to a much broader audience from many fields, including physics, computer science, mathematics, statistics, and the social sciences. The book has been the standard text for the majority of information theory courses worldwide, helping educate tens of thousands of students and attracting generations of new researchers to the field, even as the interest in information theory has expanded from communication to other applications. As a testament to the book's impact on research, it has garnered more than 61 500 citations (according to Google scholar), and this number is an underestimate of its broader impact on education.

The philosophy of the textbook is succinctly stated in the preface using a quote from Einstein "Everything should be made as simple as possible, but no simpler." This exposition enabled it to explain the fundamental mathematics of information theory in an accessible manner, without compromising on precision. Another unique aspect of the book is to make connections of information theory to several areas of science and engineering, demonstrating its importance beyond its origins. Shannon's seminal article [3] focused on a mathematical theory of communications, and his development of information theory was inspired by this. However, the book shows that information theory has fundamental impact in fields as diverse as machine learning, theoretical computer science, statistical physics, mathematics, probability and statistics, economics, etc. This is illustrated in Figure 2 adapted from Chapter 1 of the textbook. In fact, the applications of information theory to various fields have only grown since the textbook's publication, and in several of these fields the book is cited, demonstrating its impact to these areas.

Figure 1
Joy Thomas.

Joy's technical contributions, however, went far beyond his textbook. In the sections that follow we describe some of these contributions, viewed through the lens of the stages of his career. Through his unique career path, Joy demonstrated that studying information theory does not constrain one to just an academic

2692-4080 © 2021 IEEE
Digital Object Identifier 10.1109/MBITS.2021.3108946
Date of publication 1 Sept. 2021; date of current version 4 Feb. 2022.

career path and that one can use "information-theoretic thinking" to make an impact to diverse applications. As we describe next, Joy worked on a variety of problems in industry and startups, bringing to it insights and ideas from information theory, while achieving significant success. We believe that this illustrates the enduring importance of information theory to both basic science and engineering applications.

Technical Contributions

Stanford 1984–1990: Joy came to Stanford after completing his Bachelor of Technology degree at the Indian Institute of Technology (IIT), Madras, India, in 1984. He was admitted to the IIT as a 16-year old, after securing All-India Rank #1 in the famed Joint Entrance Exam (JEE) for the IITs. He joined Stanford University on the Charles LeGeyt Fortescue scholarship, awarded by the IEEE to a single freshman graduate student in electrical engineering each year. He was later awarded an IBM Ph.D. fellowship during the final three years of his Ph.D.

Joy's Ph.D. thesis, supervised by the famed information theorist Thomas Cover, contains what is perhaps the first thorough investigation on the role of feedback in networks [4]. It had already been known for several decades that feedback does not increase the capacity of memoryless point-to-point channels, though it could help simplify coding schemes. There had also been several works in the 1970s and 1980s, which showed that feedback can increase the capacity of certain networks. Joy's thesis was a tour-de-force in understanding the role of feedback in many types of networks, and at the time of its writing it had the best results for the capacity of feedback networks. It showed that feedback could increase capacity by at most a factor of two in multiple access networks (think wireless uplink, where several users communicate with a base station) [5]. It also showed that feedback can help significantly in interference networks (think wireless communication with multiple users simultaneously transmitting to different receivers). His thesis also established a set of fundamental entropic and convex-geometric inequalities, which still stand today in terms of their use and the elegance of their information-theoretic proofs, and which demonstrate the utility of information theory in proving basic inequalities in mathematics (also see Dembo *et al.*'s work [6]).

IBM Research 1990–1999: After completing his Ph.D. in 1990, Joy joined the IBM Thomas J. Watson Research Center. He worked on a diverse set of topics including data compression, connections between queueing theory and information theory, as well as the (then) emerging area of data mining. He was an early advocate for the application of information-theoretic ideas and techniques to problems involving the understanding of large-scale data, a path that led to his successful entrepreneurial efforts described later.

Along with Cheng-Shang Chang, Joy explored the interaction between information theory and queueing theory. Their collaboration started with the observation that the large deviation

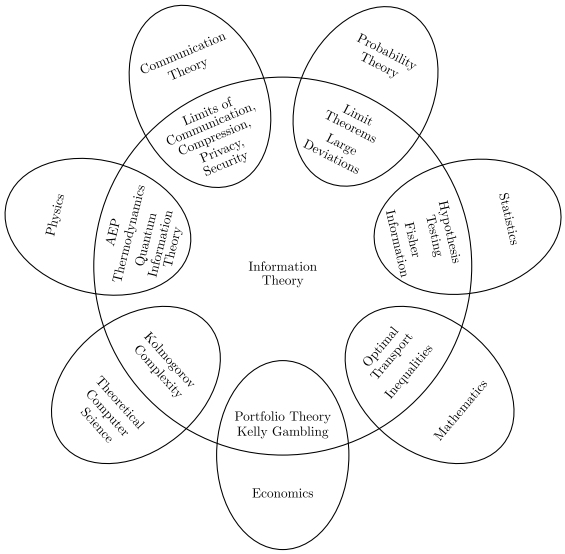

Figure 2

Connections between information theory and other fields, as demonstrated in his works [1], [2].

principle is a key mathematical tool linking these two areas. They identified the entropy functional as a key notion in describing effective bandwidth in queueing networks, and their collaboration led to two articles [7], [8] on this topic.

Joy also worked on using data compression techniques for caching in computer systems. Part of that research contributed to the development of Memory eXtension Technology (MXT), which was implemented in the IBM Pinnacle memory controller chip. That technology has lived on through numerous generations of refinement, and it is now part of IBM's Active Memory Expansion technology used across all of their computer systems, from industry standard x86 servers through their Power Enterprise servers.

The MXT idea was to insert a compressor/decompressor between cache and memory: Decompress on cache misses, compress on writeback. One needed something akin to a universal source coding algorithm (they chose Lempel–Ziv) to make this work, but Lempel–Ziv methods require a long buffer in order to be effective, and compressing buffers of that size was challenging to do quickly, even in hardware. The solution was a shared dictionary approach: Parallelize the process, allowing each compressor to share its history with the others. The results were impressive: Memory was compressed by a factor of 3, with negligible performance impact. This enabled executing workloads, that previously needed hours, to be completed in minutes. Though Joy did not write technical papers on this work, some of his contributions are documented in several patents [9]–[11].

Startups 1999–2020: In 1999, Joy left IBM to join the founding team at Stratify, a pioneer in unstructured data management, autogenerating taxonomies, and conceptual classification models (see the Thomas *et al.*'s work [12]). Joy served as Stratify's Chief Scientist, until it was acquired by Iron Mountain in 2007.

In 2011, Joy co-founded InsightsOne. At the time, due to the availability of inexpensive large-scale computing and large amounts of data, significant efforts were being made to develop predictive models for engagement and advertising using online behavioral observations. InsightsOne harnessed the commercial opportunity to use Big Data and associated tools for the purpose of designing and creating a predictive analytics platform, which could be used by the healthcare, retail, telecom, and financial services industries.

Joy was instrumental in developing technical solutions to problems arising in this platform's design. He formulated the entire project as a graph problem, where actions by agents were represented as sequences of time-stamped typed events, with arbitrary attribute-value pairs. Attributes' values could be arranged into hierarchies, giving coarser and finer attribute descriptions. The main idea behind this formulation was to facilitate the identification of patterns at differing levels of abstraction. The sequence of events for each agent could be combined into a single graph, with each node representing an action at a certain chosen level of abstraction.

Joy's key contribution was something that might seem obvious in hindsight, but was not obvious before—a characteristic he shared with his Ph.D. advisor Thomas Cover. His insight was that a sequence of interactions that a person has with a business is an excellent predictor of her behavior over the course of the next few interactions. This insight upended the state of the art at that time, which viewed interactions as a bag rather than a sequence. Joy worked out algorithms that were effective in finding predictive patterns in these high-dimensional spaces. These algorithms created a significant lift in performance across applications. This work led to several patents, including the one by Jhingran *et al* [13].

InsightsOne was acquired by Apigee in 2014, which was later acquired by Google in 2016. One of the challenges at Apigee was that decisions had to be made very quickly at "line speed." In order to protect Apigee customers from unwelcome bot attacks, Joy was able to design a collection of local "rules," each of which was simple and amenable to efficient implementation, but which when combined and applied collectively were able to get close to the quality that a much more global (and expensive) system could achieve. This balance came from his deep understanding of both the theory and the practical constraints within which the theory was to be applied.

Joy Thomas Foundation and The IEEE Information Theory Society

The Foundation: The Joy Thomas Foundation was started in 2020 by several of Joy's friends and colleagues to honor his legacy. Its goal is to inspire excellence among the next generation of STEM students, researchers, and entrepreneurs. Their network includes industry executives and researchers, faculty and college students from major global educational institutions (the Indian Institutes of Technology, Stanford University, MIT), industry research labs (IBM, Microsoft), silicon valley and India-based tech startups and large enterprises (Google, Microsoft, Cisco, VMWare, Zoho), and high-school students. The envisaged programs include (i) *Awards and Scholarships:* supporting excellence in research and education through financial awards; (ii) *STEM for college students:* Support talented college students from traditionally disadvantaged groups with scholarships, soft skills training, and mentorship to achieve STEM careers; (iii) *STEM for high-school students:* Develop mentorship programs to encourage high-school students to take the STEM path. These will be funded through both corporate and individual contributions to the Foundation.

Connections to IEEE IT Society: Since Joy's technical home was the Information Theory Society, the Foundation has pledged to support the society in several ways. The first initiative is to fund an IEEE Information Theory Society Tutorial Paper Award. This award has been approved by the IEEE Information Theory Society Board of Governors and details of the award are currently being formulated along with the IEEE. Given Joy's work as well as his talent in exposition, the Joy Thomas Tutorial Paper Award would be a fitting legacy to him as well as support the IEEE IT society to expose its research broadly, within and outside the community.

Acknowledgments

The authors gratefully acknowledge inputs from Dr. Vittorio Castelli, Prof. Cheng-Shang Chang, Dr. Waqar Hasan, and Dr. Anant Jhingran for this article.

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1991.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006. [Online]. Available: http://www.elementsofinformationtheory.com/

[3] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[4] J. A. Thomas, "The role of feedback in multiuser information theory," Ph.D. dissertation, Elect. Eng. Dept., Stanford, CA, USA, Stanford Univ., 1990.

[5] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 5, pp. 711–716, Sep. 1987, doi: 10.1109/TIT.1987.1057341.

[6] A. Dembo, T. M. Cover, and J. A. Thomas, "Information theoretic inequalities," *IEEE Trans. Inf. Theory*, vol. 37, no. 6, pp. 1501–1518, Nov. 1991, doi: 10.1109/18.104312.

[7] C. Chang, J. A. Thomas, and S.-H. Kiang, "On the stability of open networks: A unified approach by stochastic dominance," *Queueing Syst. Theory Appl.*, vol. 15, no. 1–4, pp. 239–260, 1994.

[8] C. Chang and J. A. Thomas, "Effective bandwidth in high-speed digital networks," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 1091–1100, Aug. 1995, doi: 10.1109/49.400664.

[9] P. Franaszek, J. Robinson, and J. Thomas, "Adaptive multiple dictionary data compression," U.S. Patent 5870036, Feb. 9, 1999.

[10] P. Franaszek, J. Robinson, and J. Thomas, "Parallel compression and decompression using a cooperative dictionary," U.S. Patent 5729228, Mar. 17, 1998.

[11] V. Castelli, P. Franaszek, and J. Thomas, "Method and apparatus for prediction of computer system performance based on types and numbers of active devices," U.S. Patent 8234229, Jul. 31, 2012.

[12] J. Thomas, M. Lakhamraju, G. Mathew, P. Nayak, G. Ramana, and J. O. Lamping, "Techniques for organizing data to support efficient review and analysis," U.S. Patent 7945600, May 17, 2011.

[13] A. Jhingran, K. Kesavan, J. Thomas, J. Chand, and S. Rajagopalan, "Automatically extracting profile feature attribute data from event data," U.S. Patent 10255300, Apr. 9, 2019.