Newsjournal of the Society for Industrial and Applied Mathematics sinews.siam.org

siam news

Volume 53/ Issue 8
October 2020

The Connection Between Applied Mathematics and Deep Learning

By Manuchehr Aminian

In recent years, deep learning (DL) has inspired a myriad of advances within the scientific computing community. This subset of artificial intelligence relies on multiple components of applied mathematics, but what type of relationship do applied mathematicians have with DL? This question was the subject of a plenary talk by Yann LeCun (Facebook and New York University) at the virtual 2020 SIAM Conference on Mathematics of Data Science,[1] which took place earlier this year. LeCun provided a brief history of machine learning (ML), highlighted the mathematical underpinnings of the field, presented both his vision and several broad open questions for ML's future, and discussed applied math's current relation and potential impending contributions. A 2018 *SIAM News* article by Gilbert Strang, entitled "The Functions of Deep Learning,"[2] offers an introduction for those who are unfamiliar with neural networks, ML, and DL.

Stochastic Gradient Descent

LeCun immediately identified applied mathematicians' most fundamental connection to DL: gradient descent and optimization. The search for an optimal set of parameters for a nonlinear function—with the goal of succeeding in a practical task, such as classifying images or predicting text—comprises the heart of DL. Researchers use a special form of gradient descent to find this optimal set of parameters.

Applied mathematicians often encounter gradient descent in numerical linear algebra when they seek approximate solutions of a square linear system of equations $Ax=b$. Finding a solution x^* is equivalent to finding a minimum of the function $||Ax-b||_2^2$ over all choices of x, beginning with some initial guess x_0. Researchers understand gradient descent as a process wherein they "walk down the mountain" by going in the steepest direction at every step. These actions produce a sequence of approximations x_i that is guaranteed to converge to a unique minimum for symmetric positive definite A. This is admittedly a very special class of functions that allows users to take the theory quite far, which is why it is taught in the classroom.

In contrast, the functions that one must minimize in DL—known as "loss functions"—are typically nonlinear and nonconvex, which makes theoretical guarantees much more challenging. Nevertheless, practitioners utilize gradient-based approaches and typically employ a modified version called stochastic gradient descent (see Figure 1). This stochastic component relates to the loss function's evaluation; rather than using all training data to evaluate the loss, one uses a randomly selected subset of data on each iteration of gradient descent. LeCun refers to this as "walking down the mountain in a fog," wherein each sample provides a noisy estimate of the direction. This stochastic component has found enormous practical success. "Nobody even considers anything else," LeCun said. However, opportunities still exist for researchers to provide theory that explains this success.

See **Deep Learning** *on page 4*

[1] https://www.siam.org/conferences/cm/conference/mds20

[2] https://sinews.siam.org/Details-Page/the-functions-of-deep-learning

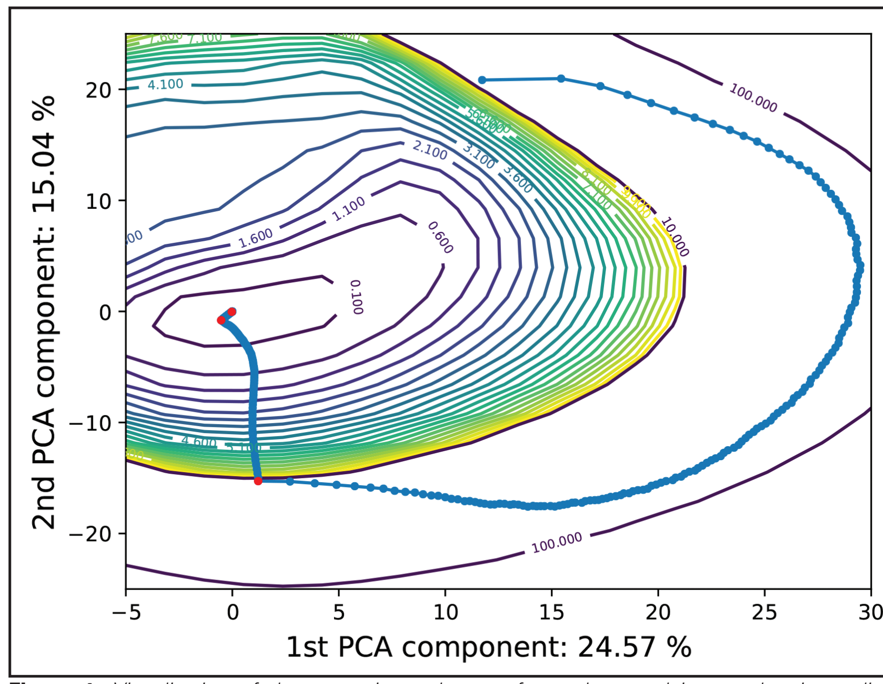


Figure 1. *Visualization of the approximate loss surface when applying stochastic gradient descent to an image dataset. Figure courtesy of Tom Goldstein.*

The Election Interference Game

By Jenny Morber

Election interference, in which one country conspires to elect a favored candidate in another country, is a classic example of a non-cooperative game. Game theory pits two opponents against one another to achieve a goal, wherein one player's path to success is contingent upon the strategy of the other. Therefore, when news of suspected Russian meddling in the 2016 U.S. presidential election made headlines, economist and mathematician David Dewhurst set out to model the allegations as optimal game play.

Critics of game theory note that it only works when both sides act rationally, whereas people often do not. Dewhurst acknowledges this incongruity but mentions an additional consideration. "If you ever wanted to look for something that is a rational actor when taken as a whole, a foreign or domestic intelligence agency is probably that," he said. "If any group has the job of acting rationally, it is them."

In a paper published earlier this year in *Physical Review E*, Dewhurst and collaborators Christopher Danforth and Peter Dodds model a two-player game in which one country, designated as "Red," wishes to influence the outcome of a two-candidate election in another country, designated as "Blue" [1]. Red's goal is the election of preferred candidate A, whereas Blue's goal is an election that is free from interference — a win that is much more difficult to define. These uneven objectives put the defensive country at a disadvantage. "If you're the FBI, you don't want Hillary Clinton to win and you don't want Donald Trump to win," Dewhurst said of the 2016 election. "You just want a free and fair election. The issue is that if you take that strategy, Red always wins. So if you want to stop Red from interfering, you actually have to interfere on behalf of one of the other candidates."

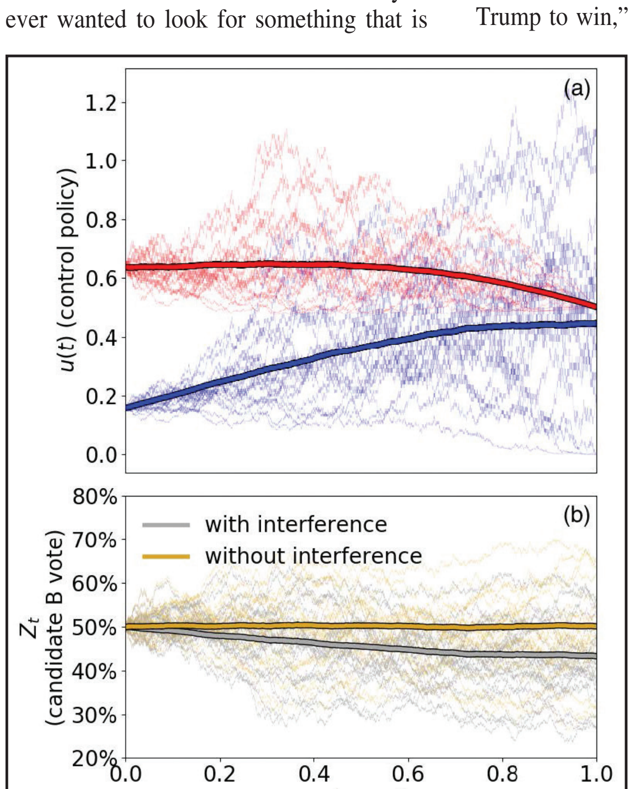


Figure 1. *Examples of control policies for Red* (u_R) *and Blue* (u_B) *in a simulated election game.* **1a.** *Optimal expenditures by Red and Blue.* **1b.** *Election results over time. Even when Blue plays optimally to resist, Red is able to influence election results and hinder candidate B. Image courtesy of [1].*

When crafting their model, the researchers considered an election between two candidates, denoted A and B, that was decided by majority vote with no Electoral College. They assume that a public poll $Z_t\in[0,1]$ represents the election process at any time $t\in[0,T]$. The model's dynamics occur in a latent space that is related to the polling process $Z_t=\phi(X_t)$, where ϕ is the sigmoidal function $\phi(x)=\dfrac{1}{1+e^{-x}}$. In this space, $X_t<0$ represents values that favor candidate A and $X_t>0$ represents values that favor candidate B.

The functions by which Red and Blue attempt to influence or deflect influence on the election are one-dimensional continuous-time stochastic processes, denoted by $u_R(t)$ and $u_B(t)$. Dewhurst and his team interpret these "control policies" as expenditures on interference operations. Under the influence of both countries' policies, the election dynamics become

$$\mathrm{d}X_t=F(u_R(t),u_B(t))\,\mathrm{d}t+\sigma\,\mathrm{d}\mathcal{W}_t,\qquad X_0=y. \tag{1}$$

Based on the assumption that F is at least twice continuously differentiable, the researchers approximate the state equation as

$$\mathrm{d}X_t=[u_R(t)+u_B(t)]\mathrm{d}t+\sigma\mathrm{d}\mathcal{W}_t,\qquad X_0=y. \tag{2}$$

Red and Blue then seek to minimize the cost functions of their own control policies. C_R and C_B respectively represent the running cost or benefit of conducting election interference operations, according to

$$E_{u_R,u_B,X}\left\{\Phi_R(X_T)+\int_0^T C_R(u_R(t),u_B(t))\mathrm{d}t\right\} \tag{3}$$

and

$$E_{u_R,u_B,X}\left\{\Phi_B(X_T)+\int_0^T C_B(u_R(t),u_B(t))\mathrm{d}t\right\}. \tag{4}$$

Here, the cost functions take the form $C_i(u_R,u_B)=u_i^2-\lambda_i u_{\neg i}^2$ for $i\in\{R,B\}$. The notation $\neg i$ indicates the set of all other players. Therefore, if $i=R$, $\neg i=B$ and λ_i parameterizes the utility that player i

See **Election Interference** *on page 2*

Nonprofit Org
U.S. Postage
PAID
Permit No 360
Bellmawr, NJ

siam®
SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

siam news
Volume 53/ Issue 8/ October 2020

Election Interference

Continued from page 1

gains by observing player $\neg i$'s effort. If $\lambda_i > 0$, player i gains utility when player $\neg i$ expends resources.

To find optimal play, Dewhurst and his colleagues define final conditions that specify the costs that Red and Blue incur, then solve the game backward through time. Because Red wants to influence the election's outcome in favor of candidate A, its final cost function Φ_R must satisfy $\Phi_R(x) < \Phi_R(y)$ for all $x < 0$ and $y > 0$. The researchers considered three possibilities for Blue's final conditions, as they are much more complex. One possible condition is that Blue may accept the election result if it does not deviate "too far" from the initial expected value. A discontinuous condition that represents this preference is given by $\Phi_B(x) = \Theta(|x| - \Delta) - \Theta(\Delta - |x|)$, where $\Delta > 0$ is Blue's accepted margin of error and $\Theta(\cdot)$ is the Heaviside step function.

Applying the dynamic programming principle to equations (2)-(4) yields the following system of coupled Hamilton-Jacobi-Bellman equations for the Red and Blue value functions:

$$-\frac{\partial V_R}{\partial t} = \min_{u_R} \left\{ \frac{\partial V_R}{\partial x}[u_R + u_B] + u_R^2 - \lambda_R u_B^2 + \frac{\sigma^2}{2}\frac{\partial V_R}{\partial x^2} \right\} \tag{5}$$

and

$$-\frac{\partial V_B}{\partial t} = \min_{u_B} \left\{ \frac{\partial V_B}{\partial x}[u_R + u_B] + u_B^2 - \lambda_B u_R^2 + \frac{\sigma^2}{2}\frac{\partial V_B}{\partial x^2} \right\}. \tag{6}$$

Minimizing these equations with respect to the control variables then produces the Nash equilibrium control policies:

$$u_R(t) = -\frac{1}{2}\frac{\partial V_R}{\partial x}\bigg|_{(t,X_t)} \tag{7}$$

and

$$u_B(t) = -\frac{1}{2}\frac{\partial V_B}{\partial x}\bigg|_{(t,X_t)}, \tag{8}$$

as well as the exact functional forms of equations (5) and (6). When solved over the entire state space, the latter become the strategies of a subgame perfect Nash equilibrium:

$$-\frac{\partial V_R}{\partial t} = -\frac{1}{4}\left(\frac{\partial V_R}{\partial x}\right)^2 - \frac{1}{2}\frac{\partial V_R}{\partial x}\frac{\partial V_B}{\partial x} - \frac{\lambda_R}{4}\left(\frac{\partial V_B}{\partial x}\right)^2 + \frac{\sigma^2}{2}\frac{\partial^2 V_R}{\partial x^2},$$
$$V_R(x,T) = \Phi_R(x) \tag{9}$$

and

$$-\frac{\partial V_B}{\partial t} = -\frac{1}{4}\left(\frac{\partial V_B}{\partial x}\right)^2 - \frac{1}{2}\frac{\partial V_B}{\partial x}\frac{\partial V_R}{\partial x} - \frac{\lambda_B}{4}\left(\frac{\partial V_R}{\partial x}\right)^2 + \frac{\sigma^2}{2}\frac{\partial^2 V_B}{\partial x^2},$$
$$V_B(x,T) = \Phi_B(x). \tag{10}$$

Dewhurst and his collaborators employ backward iteration to numerically identify the value functions $V_R(x,t)$ and $V_B(x,t)$. This enforces a Neumann boundary condition at $x = \pm 3$ that corresponds to a bound on the polling popularity of candidate B by 4.7 percent from below and 95.3 percent from above.

Equations (5) and (6) provide the closed-loop control policies for Red (u_R) and Blue (u_B), given the current state X_t and time t. Figure 1 (on page 1) displays examples of u_R, u_B, and the electoral process Z_t. In this example, the researchers simulate the game with parameters $\lambda_R = \lambda_B = 2$, $\Phi_R(x) = x$, and $\Phi_B(x) = \frac{1}{2}x^2\Theta(-x)$. The control policies—the amount of resources spent attempting to win the election—are plotted in Figure 1a (on page 1); the thick curves illustrate the average values. Figure 1b shows the path of the electoral process. Red's optimal play involves beginning with a large amount of interference and decreasing interference on average over time. Blue is at a clear disadvantage, as optimal play by both players results in lower vote counts for candidate B than the electoral process with no interference. Even though Blue resists Red's interference, candidate A wins the election and Red still accomplishes its objective. In many simulations, Red can interfere even when Blue plays optimally. This analysis suggests that Blue must invest resources in electing candidate B to achieve results that imitate a free and fair election.

Dewhurst's team then conducts a coarse parameter sweep over λ_R, λ_B, Φ_R, and Φ_B to explore the game's qualitative behavior. Holding Blue's final condition of $\Phi_B(x) = \frac{1}{2}x^2\Theta(-x)$ constant, they compare the means and standard deviations of the Nash equilibrium strategies $u_R(t)$ and $u_B(t)$ across values of the coupling parameters $\lambda_R, \lambda_B \in [0,3]$ as Red's final condition changes from $\Phi_R(x) = \tanh(x)$ to $\Phi_R(x) = \Theta(x) - \Theta(-x)$.

Some combination of these parameters leads to an "arms race," in which Nash equilibrium strategies inspire super-exponential growth in both players' control policies towards the end of the game. In short, an all-or-nothing mindset causes Red and Blue to spend large amounts of money for no change in outcome. Therefore, Blue may choose to let the attacking country interfere a bit to avoid an arms race and gain additional intelligence benefits. For Red, a partial win might tighten the election and make candidate B feel less surefooted and supported.

In 2015 and 2016, Russian military foreign intelligence sought to interfere in the U.S. presidential election in favor of Donald Trump. Evidence shows that they worked in part through social media platforms like Facebook and Twitter. When this attack was discovered, Twitter shut down accounts associated with Russian government (Red team) activity and collected and analyzed the corresponding data.

To test their model, Dewhurst and his colleagues tapped this publicly available data. Since they could not observe Red or Blue's control policies, they used tweets sent by Twitter accounts associated with Russian military intelligence as proxy. The group downloaded nearly three million tweets from 2,848 unique Twitter handles, which were collected by Darren Linvill and Patrick Warren of Clemson University and hosted by FiveThirtyEight[1] [2]. 1,107,361 of these tweets were sent in the year before the 2016 election. The researchers grouped the time series of total tweet numbers by day to infer u_R and utilized RealClearPolitics poll data to mimic the electoral process, ignoring the minor effects of other parties. Using this information, they identified model parameters that best explained these inferred control policies and ultimately demonstrated that their model offers a sound explanation of the inferred variables.

This work represents a simplified construct of a complex and messy system and therefore contains inherent limitations. For example, it does not account for the U.S. Electoral College or minor political parties. The random walk assumption ignores the deterministic influence of past polls on future choices. No strategies account for other major news events, and the model does not consider tactics that influence non-voting. Yet despite its simplicity, the model provides valuable qualitative insight into U.S. and Russian strategies regarding the 2016 presidential election and may be informative for future two-candidate political races.

References

[1] Dewhurst, D.R., Danforth, C.M., & Dodds, P.S. (2020). Noncooperative dynamics in election interference. *Phys. Rev. E, 101*, 022307.

[2] Linvill, D.L., & Warren, P.L. (2018). Troll factories: The internet research agency and state-sponsored agenda building. *Resource Centre on Media Freedom in Europe*. Retrieved from http://pwarren.people.clemson.edu/Linvill_Warren_TrollFactory.pdf.

Jenny Morber holds a B.S. and Ph.D. in materials science and engineering from the Georgia Institute of Technology. She is a professional freelance science writer and journalist based out of the Pacific Northwest.

[1] https://github.com/fivethirtyeight/russian-troll-tweets/

SIAM Welcomes Suzanne L. Weekes As New Executive Director

In September, SIAM announced the appointment of Suzanne L. Weekes (Worcester Polytechnic Institute) as the new executive director, starting January 1, 2021. Weekes is replacing Jim Crowley, who retired last month after nearly 26 years at the helm.

Weekes holds a B.S. in mathematics from Indiana University and a Ph.D. in mathematics and scientific computing from the University of Michigan. She currently serves on SIAM's Council and is a member of the SIAM Committee on Science Policy, among multiple other involvements and achievements. Read more about Weekes' experience in the full announcement, which is available on SIAM News Online.[1]

We look forward to her leadership and extend our thanks to the search committee for its diligent efforts.

[1] https://sinews.siam.org/Details-Page/dr-suzanne-l-weekes-named-siam-executive-director

siam news

ISSN 1557–9573.

Update from the 2020 SIAM Council Meeting

By James Crowley

As with most gatherings during this time, the annual SIAM Council meeting commenced virtually this summer. The Council met in two sessions to allow for time zone differences. In addition to the usual matters, it addressed a significant number of new business items, including reports from all of the SIAM officers and the executive director. Throughout the course of routine procedure, the Council offered suggestions for 2021 membership dues and recommended that the SIAM Board of Trustees approve all 14 activity group (SIAG) charter renewals. It deferred discussions on pricing the 2021 SIAM Annual Meeting until further information—such as the cost of a possible hybrid meeting, which would accommodate both in-person and remote attendees—is available.

During the meeting, the Council voted in favor of offering an additional two-year term to two vice presidents (VPs): VP for Education Katie Kavanagh (Clarkson University) and VP for Science Policy Anne Gelb (Dartmouth College). The group also approved Sharon Arroyo (Boeing Company) as VP for Industry, replacing retiring VP Amr El-Bakry (ExxonMobil). VPs are appointed by the SIAM president with the advice and consent of the Board and Council. The exception is the VP-at-Large, who is elected by SIAM membership.

The Council also elected Talitha Washington (Atlanta University Center Consortium Data Science Initiative) to serve on the SIAM Nominating Committee. This committee selects the slate of candidates for the fall elections of elected officers and Board and Council members.

One SIAG charter renewal featured an unprecedented twist. Upon request from its officers, the SIAG on Data Mining and Analytics[1] became the more general SIAG on Data Science (see sidebar). The group will continue to cover data mining and sponsor the SIAM International Conference on Data Mining, but it will now also organize the SIAM Conference on Mathematics of Data Science, which held its inaugural meeting[2] online earlier this year.

Some major new business items are as follows:

– Creating a new position: VP for Equity, Diversity, and Inclusion (EDI)

– Voting to approve Ron Buckmire (Occidental College) as the new VP for EDI

– Establishing a Policy for Honors and Awards

– Establishing a Code of Conduct for SIAM members

– Voting to make the ad hoc Committee on Ethics a standing committee

– Establishing disclosure forms for honors, awards, and SIAM leadership—to be completed by nominators for prizes and SIAM Fellows, prize and Fellow recipients, and individuals running for SIAM offices—to certify that no known ethical issues exist

– Recommending to the Board that SIAM become a Level A partner in the newly-revised Joint Mathematics Meetings, beginning in 2022

– Creating a Prize Canvassing Committee

– Approving a new (revised) policy on potentially offensive materials for SIAM journals and conferences, which specifically bans the Lena image

– Reviewing a significant report from the Major Awards Committee, chaired by VP-at-Large Carol Woodward (Lawrence Livermore National Laboratory), which reviewed the SIAM Prize Program.

[1] https://www.siam.org/membership/activity-groups/detail/data-mining-and-analytics

[2] https://www.siam.org/conferences/cm/conference/mds20

VP for EDI

Creation of the VP for EDI position guarantees the presence of a voice at the Cabinet level that promotes and ensures equity and inclusion in all facets of SIAM. The VP for EDI will be a member of the SIAM Council who establishes SIAM policy and offers input on Board discussions. The position will also oversee the Membership Committee (which previously reported to the VP-at-Large) and the Diversity Advisory Committee.

Policy for Honors and Awards

SIAM joined over 130 scientific and engineering societies in the Societies Consortium on Sexual Harassment in STEMM.[3] The Societies Consortium produced a model policy, which SIAM adopted with minor modifications. When SIAM confers an honor or award upon an individual, this action reflects SIAM's judgment that the individual's contributions to—and effects on—the field of applied mathematics are exemplary. Unethical conduct by a current or prospective awardee can contribute to longstanding and systemic barriers in the field. Therefore, in order to more heavily emphasize what is best for excellence in the field rather than a single individual, SIAM will not confer an honor upon any person who has engaged in unethical conduct, or about whom a credible but undetermined question exists. The policy also allows for multiple honors, such as Fellows status, to be revoked with just cause.

Code of Conduct

Because the Policy for Honors and Awards refers to a standard of ethics, the Committee on Ethics—convened by SIAM President Lisa Fauci (Tulane University)—proposed a Code of Conduct that sets expectations for professional behavior related to harassment and issues of scientific misconduct. The Council modified and approved these guidelines, which will soon be posted online.

Disclosure Forms

It is important that SIAM promote inclusion and foster a culture of ethical behavior. While this has generally been the case in the SIAM community, the Committee on Ethics felt that a disclosure form would further bolster the importance of ethical behavior in professional life. As such, candidates for leadership positions and nominees for SIAM prizes and honors must disclose any past allegations or institutional proceedings that resulted in a finding of professional misconduct, as well as any current formal complaints that are related to professional conduct — even if the matter is still pending. Nominators must also certify that they are not aware of any ethical issues on the part of the nominee. These procedures will apply to prizes, SIAM Fellows, and nominations for leadership positions at SIAM.

Prize Canvassing Committee

For several years, SIAM's Fellows Canvassing Committee has encouraged the nomination of individuals from underrepresented groups for the Fellows Program. Such groups include SIAM members who work in industry, members outside of North America, and members from underrepresented demographics with respect to gender, race, and ethnicity. Canvassing committees are independent from selection committees and thus have no role in the ultimate selections. Like the Fellows Canvassing Committee, the new Prize Canvassing Committee will encourage a diverse nomination pool.

[3] https://sinews.siam.org/Details-Page/siam-joins-societies-consortium-on-sexual-harassment-in-stemm

New SIAM Activity Group on Data Science

At the July meeting of the SIAM Council, the Council approved the SIAM Activity Group on Data Science (SIAG/DATA) as a broader rescope of the SIAM Activity Group on Data Mining and Analytics (SIAG/DMA), which was founded in 2011. The new SIAG/DATA will aim to advance the mathematics of data science. As such, it will focus on the mathematical, statistical, and computational foundations of data science, in addition to data science's applications to other fields of science and across technology and society. SIAG/DATA will augment SIAM's *Journal on Mathematics of Data Science* (*SIMODS*), which published its first batch of articles in 2019, by providing support for this field within SIAM. The recently initiated biennial SIAM Conference on Mathematics of Data Science (MDS) will serve as the SIAG's flagship conference. The inaugural MDS20 took place virtually this May, and the next installment is scheduled for May 2022.

SIAG/DATA will continue to sponsor the annual SIAM International Conference on Data Mining (SDM). The initial slate of officers for the SIAG are Hans De Sterck (chair), Gitta Kutyniok (vice chair), Karen Willcox (program director), and Danai Koutra (secretary); they are appointed for the duration of one year (2021). In the second half of 2021, members of the SIAG will elect new officers for the two-year period of 2022-2023. SIAM members with a research interest in data science will have the opportunity to join SIAG/DATA as part of their SIAM membership renewal in the fall of 2020.

— Hans De Sterck, University of Waterloo

Revised Policy of Potentially Offensive Materials

SIAM journals have had a policy in place regarding potentially offensive materials for many years. However, authors continue to submit papers or books with images of Lena, which is considered particularly offensive to many members of the SIAM community. Objection to this image has long been publicly expressed. Until now, SIAM conferences had no written policy about offensive materials. At its most recent meeting, the Council approved the adoption of a policy that explicitly bans the Lena image—along with any other offensive materials—for both journals *and* conferences.

Major Awards Committee

Finally, the Major Awards Committee—which is chaired by the VP-at-Large and oversees all SIAM prizes—conducted a major review of the entire SIAM Prize Program, including both SIAM-wide honors and SIAG awards.

As evidenced by the numerous outcomes, this was an extremely busy and productive session for the SIAM Council. We offer our sincere thanks to Council members—especially SIAM President Lisa Fauci, who chairs the Council—for their hard work and diligence under the duress of the ongoing pandemic.

James Crowley retired in September after serving as the executive director of SIAM for nearly 26 years.

Deep Learning

Continued from page 1

Overparameterization and Deep Nets

Turning to areas that are open to theoretical discovery, LeCun first highlighted a phenomenon that is contrary to traditional mathematical and statistical intuition: understanding overparameterized models in DL. Mathematician John von Neumann famously said, "*With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.*" This sentiment reflects a common attitude among mathematical modelers and suggests that one should be mindful of both the number of parameters in a mathematical model and the conclusions that result from fitting the model to data. Mathematicians are familiar with the perils of overparameterization from fitting high-degree polynomials. Similarly, parameter fitting and identifiability are common issues in differential equation models, especially when these models have many interacting components.

However, years of practice in DL reveal a different picture. Consider two neural networks with the same goal. The network with *more* parameters—often with orders of magnitude more parameters than data points—will numerically converge in the loss function, fit the training data, and still successfully predict unseen data. In contrast, the "smaller" net frequently gets trapped in local minima in the loss surface and struggles to converge. LeCun indicated that researchers commonly understand that these overparameterized nets will automatically reduce their "rank" during training by implicit regularization, but admitted that this aspect is still a theoretical mystery. He thus suggested that the applied math community could perhaps contribute to the knowledge surrounding this problem.

The Value of an Applied Mathematics Background

DL has found particular success in the area of image processing via convolutional neural nets (CNNs). LeCun provided many examples of these triumphs during his talk, including medical image analysis, self-driving vehicles, and automated emergency braking systems. However, convolutional approaches in DL succeed with image-based applications because images are rectangular lattices; they face challenges when generalizing to other graph-based applications that lack this structure.

A generalization of the convolutional net to an arbitrary graph is called a graph convolutional net. LeCun understands the extension of CNN tools to arbitrary graphs in terms of Fourier transforms. To employ the convolution on the graph, one applies Fourier transforms to the data and filter, multiplies, and then applies the inverse Fourier transform. Unlike computer scientists, most applied mathematicians are already quite familiar with these tools. They are consequently in a strong position to understand how application problems fit in a more general theoretical landscape and are thus useful in the DL community.

However, the relationship between the applied mathematics and DL communities is not one-directional. LeCun emphasized that practitioners are interested in developing DL approaches that accelerate the numerical solution of partial differential equations (PDEs). Traditional approaches to solving PDEs rely on the solution of finite difference or finite element discretizations of the differential operator. However, issues related to discretization for the time and space variables arise for stiff and/or high-dimensional problems. In areas where careful obedience of the physics makes numerical solution difficult, one may attempt to replace the solution operator with a neural net that is trained to produce a solution based on a class of examples. While accuracy, precision, and preservation of conserved quantities are issues for such neural networks, the potential speed-ups are quite promising. During his talk, LeCun alluded to applications in lattice quantum chromodynamics, fluid dynamics, and astrophysics.

A Unifying Perspective

In the past, neural networks were primarily motivated by a desire to understand the living brain. Therefore, LeCun's presentation also touched on DL approaches that mimic the ways in which humans learn, reason, and plan complex tasks. To quote LeCun, humans are "barely supervised and rarely reinforced." Why is this? Success with deep networks in image processing, for instance, requires thousands or even millions of labeled examples and an enormous amount of computational power for training purposes. The neural net may find success in the same class of images with which it was trained, but the process must begin again upon the introduction of a fresh class of images that were not seen during training. The user must also alert the machine to the new class of objects.

This is in stark contrast to the way that babies learn. They can recognize new objects after only seeing them a few times, and do so with very little effort and minimal external interaction. If ML's greatest goal is to understand how humans learn, one must emulate the speed at which they do so. This direction of research is exemplified by a variety of techniques that may or may not fit into an existing paradigm; LeCun classified these tasks under the umbrella of "self-supervised learning."

While supervised learning and reinforcement learning have shown success in isolated tasks, LeCun believes that these paradigms will never lead to so-called "artificial general intelligence," regardless of the scale-up of hardware capabilities. Throughout his talk, he alluded to some of the fundamental challenges that are associated with these approaches. In particular, LeCun feels that reinforcement learning will struggle to explore state space, especially when one imposes the "rarely reinforced" aspect of natural intelligence. He stated that reinforcement learning will thus make it difficult for researchers to even develop a system that has "cat-level intelligence," much less human-level intelligence. As an alternative, LeCun discussed self-supervised learning via energy-based models during the second half of his lecture. This approach foregoes neural nets altogether and instead learns an energy surface that captures dependencies between inputs and can allow for multiple output predictions. He briefly highlighted many approaches but predicted that regularized latent-variable energy-based models will be the winning framework.

LeCun concluded by addressing a pervading question: Is DL a natural science or an engineering science? Is it *truly* a science, or more like alchemy? He implicitly conceded that it is mostly the latter but argued that "just because we don't understand it doesn't mean we shouldn't use it." LeCun noted that the gap between discovery of a machine and the subsequent identification of its theory can historically be quite lengthy. For example, the telescope was developed in 1608, but it took 50 years for optics to explain why it works. The steam engine appeared in 1695, but more than 100 years passed before thermodynamics could describe its function. LeCun hopes to "find the equivalent of thermodynamics for machine intelligence, or intelligence in general." Mathematicians will likely play a significant role in this endeavor before the DL community can reach anything that resembles a unified theory.

This article is based on Yann LeCun's invited talk at the 2020 SIAM Conference on Mathematics of Data Science (MDS20), which occurred virtually earlier this year. LeCun's presentation is available on SIAM's YouTube Channel.[3]

Manuchehr Aminian is an assistant professor in the Department of Mathematics and Statistics at California State Polytechnic University, Pomona. His interests include mathematical modeling, visualization, and mathematical methods in data science.

[3] https://www.youtube.com/watch?v=9sqoe_krQ1E

Virtual Summer Schools: Can We Make Them Work?

By Chris Jones and Hans Kaper

Summer schools are integral parts of students' professional training in the mathematical sciences. Unfortunately, COVID-19 rendered in-person summer school experiences unfeasible this year. Having recently completed a six-week virtual summer school with 42 students who spanned four different time zones, we decided to reflect on the experience and summarize some of our takeaways.

The summer school on "Dynamics and Data in the COVID-19 Pandemic"[1] was hosted by the American Institute of Mathematics (AIM), a U.S. mathematical science research institute that is supported by the National Science Foundation. The program was organized in collaboration with the Mathematics and Climate Research Network (MCRN),[2] a virtual research network that engages mathematicians in climate research. Chris Jones (University of North Carolina at Chapel Hill) led the online summer school, which ran from June 22 to July 31, 2020.

The idea for the summer school was proposed in early April of 2020, and the first announcement published on May 31. We originally planned to invite 20 advanced undergraduate and graduate students, each of whom would receive a stipend of $6,000. The school was meant to be very interactive, with structured working days consisting of large and small group discussions, one-on-one meetings, and time for individual work. All interactions were to take place online via web-based infrastructure.

As the number of applications kept growing—eventually reaching well over 500—AIM and the organizing committee decided to double the allotted number of participants. This change required some organizational adjustments. The organizers hired several fully-paid mentors—junior faculty and advanced graduate students—to assist the faculty leaders, in addition to two student assistants to identify any emerging issues and facilitate interactions among students, mentors, and leaders. In total, 59 people participated: 42 students, five mentors, 10 faculty leaders, and two assistants. The summer school also retained eight experts in mathematical epidemiology, statistics, medicine, and the information technology health industry. We all met for at least five hours each day, and everyone was on a first-name basis.

[1] https://aimath.org/workshops/upcoming/mcrn2020

[2] http://mathclimate.net

Sococo

When organizing a virtual summer school, major challenges include building a sense of community and keeping participants engaged. An in-person environment addresses both of these issues naturally; everybody is in one place, often staying in the same building and meeting regularly, with the capacity for individual work as necessary. Video conferencing is the obvious solution, but too much screen time each day—in an attempt to replicate the "being together" of an in-person school—causes exhaustion among everyone.

Keeping these challenges in mind, we sought a platform that would allow us to be present and together without a constant Zoom meeting. Sococo offers such an environment by creating a virtual workspace that effectively simulates a real one (see Figure 1). The visual layout resembles an architect's plan without the measurements but with other adornments that make it look inviting, such as tables, chairs, and the occasional plant. The 59 participants and eight experts (each represented by an avatar) had their own private offices where they could hang out. We could leave our office doors open or closed, and anybody could come in by either clicking on the space (if open) or knocking on the door (you even heard the "tap-tap").

Figure 1. *The Sococo platform creates a virtual workspace where participants can gather in team rooms, lounges, and even a kitchen and café. Here, groups meet in private offices.*

The Sococo layout comprises meeting rooms, team rooms, lounges, and even a kitchen and café. There is a library to store references and a records room that holds documents of completed work; both are accessible from other spaces via clickable icons. We used the café to post the next day's schedule and a list of names with contact information, and the kitchen for sign-up sheets and the daily blog. We all met together two or three times each day in the large "all hands" meeting room, where the daily assignments (readings, research, or videos) were posted on a bulletin board.

Sococo has a native video conferencing capability that is sufficient for small group meetings, such as an intimate conversation in a virtual office. But larger meetings require the use of Zoom, Webex, or Google Meet, which one can integrate into Sococo. When a participant begins a meeting, a Sococo pop-up dialog box invites everyone else in the room. With only one click, participants can enter the meeting with audio and video.

We used Slack, which can also be integrated into Sococo, for online chatting. Slack is particularly beneficial for a summer school like ours because it is organized by channels. Specialized interest groups that emerged throughout the program could just start a new channel.

Unfortunately, virtual platforms cannot truly replicate chance encounters in the corridor, which are unique to in-person meetings. However, we tried to engineer these encounters by stimulating visits among the participants during times such as lunch breaks. On the other hand, the virtual world offers possibilities that are not feasible—or at least not natural—in its physical counterpart. For example, several participants posted short personal videos about themselves with clickable links that other students could access from any location.

Hardware and Software

Students were required to use a tablet (iPad, Microsoft Surface, or Samsung tablet) and a digital pen or pencil. A $1,000 supplement to each stipend covered the cost of this hardware. During working sessions, students shared virtual whiteboards on Google Jamboard to jot down and explore ideas. These interactions mimicked the experience of meeting in an office and working on a blackboard together.

Daily Program

Each day, the summer school met from 11:00 a.m. to 1:30 p.m. and again from 3:00 p.m. to 5:30 p.m. EDT. Students were required to be on Sococo during these periods — "in the office" and available for meetings, either planned or spontaneous. The meeting times accommodated students in four time zones, from California to Puerto Rico. The daily program was bracketed by "all hands" meetings at 11:00 a.m. and 5:15 p.m., and each afternoon program included a 15-minute tai-chi session to keep everybody invigorated.

Program by Week

Since the topic of the summer school was new to most participants, we split the program into two parts. Students spent the first three weeks learning and formulating questions, and the last three weeks conducting research to address these questions.

Participants identified the questions through articles (many from *SIAM News*[3]), online videos, and simulators. We employed various interactive exercises, such as role playing and anonymous online polling, to work through their suggestions. We also discovered that Watch2Gether—an app that allows multiple viewers to watch a video synchronously while stopping, restarting, speeding up, or slowing down the presentation—is a very useful platform for viewing online lectures. After quickly realizing that watching presentations with 50-60 people in one room was unwieldy and ineffective, we began assigning the videos as homework instead. Small groups of six or seven students then reviewed key points in Watch2Gether sessions the next morning (held concurrently with Jamboard

See **Summer Schools** *on page 7*

[3] https://sinews.siam.org/Details-Page/tag/covid-19

Leaders	
John Gemmer	Wake Forest University
Sarah Iams	Harvard University
Chris Jones	University of North Carolina at Chapel Hill
Hans Kaper	Georgetown University
Dick McGehee	University of Minnesota
Nancy Rodriguez	University of Colorado Boulder
Stephen Schecter	North Carolina State University
Mary Silber	University of Chicago
Erik Van Vleck	University of Kansas
Mary Lou Zeeman	Bowdoin College
Mentors	
James Broda	Bowdoin College
Punit Gandhi	Virginia Commonwealth University
Kaitlyn Martinez	Colorado School of Mines
Christian Sampson	University of North Carolina at Chapel Hill
María Sánchez-Muñiz	University of Minnesota

Figure 2. *List of leaders and mentors who helped organize and run the summer school on "Dynamics and Data in the COVID-19 Pandemic."*

SPPEXA: Software for Exascale Computing

By Severin Reiz and Hans-Joachim Bungartz

In 2011, the German Research Foundation (DFG) established a nationwide Priority Programme — a coordinated funding scheme open to research consortia from all of Germany. The resulting program, called Software for Exascale Computing (SPPEXA),[1] addresses fundamental research on various aspects of high-performance computing (HPC) software. SPPEXA involves 17 consortia, 39 research institutions, and 57 principal investigators that hail from the fields of computer science, mathematics, natural sciences, engineering sciences, and life sciences. Funding commenced in 2013, comprised two funding phases, and ended in April 2020. Overall funding from the DFG included 23 million euros.

When referring to exascale computing, one typically emphasizes the corresponding high-end computer systems for several reasons. First, the Top500 list[2] of the 500 most powerful systems worldwide—which publishes twice a year and typically culminates in a "race to exascale"—is always celebrated as a huge competition between technological approaches, vendors, and nations. Second, repeatedly providing large-scale

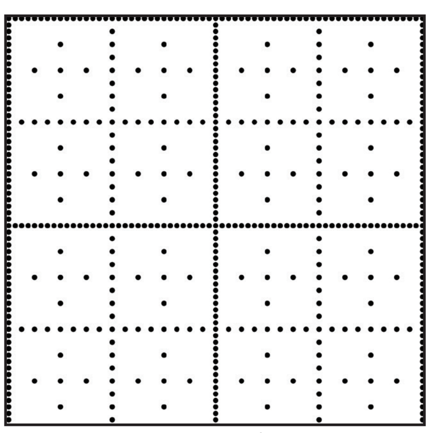


Figure 1. *Approximation of high-dimensional problems based on sparse grids. In this case, the grids are two dimensional. Image adapted from [1].*

HPC capacities is considered a feat of scientific and economic strength. And third, despite their prices (or maybe even because of them), HPC systems are frequently easier to finance as one-time investments than scientific staff. Nevertheless, one can only exploit a computer's theoretical performance in practice with efficient and scalable algorithms and a high-performance software ecosystem. Funding agencies in the U.S. and Japan quickly reacted to this challenge with their own respective programs; SPPEXA is the DFG's response.

SPPEXA provides a holistic approach to HPC software. It ensures the efficient use of current and upcoming high-end supercomputers by exploring both evolutionary and disruptive research threads. The following six research directions are considered most crucial: (i) computational algorithms, (ii) application software, (iii) system software and runtime libraries, (iv) programming, (v) software tools, and (vi) data management. Computational algorithms—such as fast linear solvers or eigensolvers—are core numerical components of many large-scale application codes, including those driven by classical simulations or oriented towards data analytics. If one cannot ensure scalability for such core routines, the battle is already nearly lost. Application software acts as the "user" of HPC systems and typically appears as legacy codes that researchers have developed over years or even decades. Increasing the software's performance via a co-design that combines algorithm and performance engineering and addresses both the "systems–algorithms" and "algorithms–applications/models" interfaces is vital.

Performance engineering cannot succeed without progress in compilers, monitoring, code optimization, verification support, and parallelization support (like auto-tuning); this dependence collectively underlines the importance of system software, runtime libraries, and tools. Programming—including programming models—is likely the research direction where the need for a balance of evolutionary research (improving and extending existing programming models, for example) and revolutionary approaches (exploring new programming models and language concepts, such as domain-specific languages) is most apparent. In addition, data management has always been relevant to HPC in terms of input/output or post-processing and visualization. It is therefore becoming increasingly important as more HPC applications are related to data.

SOFTWARE AND PROGRAMMING

To illustrate the wide spectrum of SPPEXA's research impact, we highlight three exemplary subprojects:

Algorithms: An Exa-Scalable Two-Level Sparse Grid Approach for Higher Dimensional Problems in Plasma Physics and Beyond (EXAHD)[3] supports Germany's long-standing research in the use of plasma fusion as a clean, safe, and sustainable carbon-free energy source (see Figure 1). The EXAHD initiative aims to develop scalable, efficient, and fault-tolerant algorithms to run on distributed systems, thus advancing the progress of cutting-edge plasma fusion research.

Monitoring Tools: ExtraPeak's[4] research focuses on performance optimization of parallel programs by measuring and analyzing their runtime behavior (see Figure 2). Researchers utilized *Scalasca*, a tool developed by ExtraPeak, to optimize a Barnes-Hut algorithm variant for neurons; this mechanism connected SPPEXA to the *Human Brain Project*,[5] an ongoing effort by the European Union to build a cutting-edge research infrastructure for neuroscience and computing.

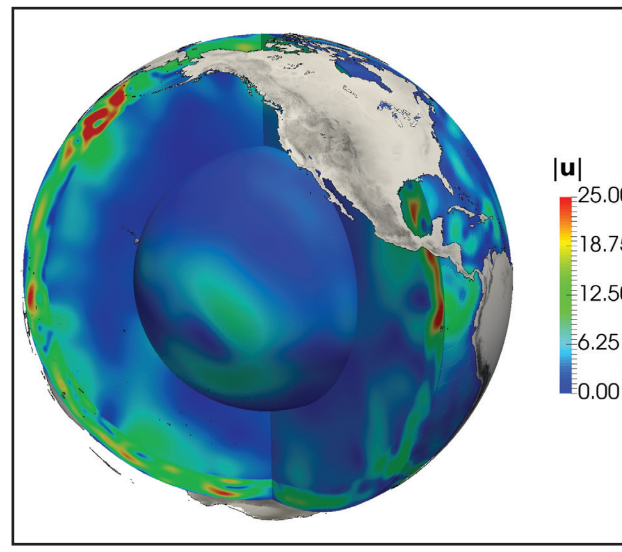


Figure 3. *Flow speeds in a mantle convection model that is computed with hierarchical hybrid grids. Image courtesy of [1].*

Applications: Terra Neo[6] addresses the challenges of understanding the convection in Earth's mantle, which is responsible for most of the planet's geological activity — from plate tectonics to volcanoes and earthquakes (see Figure 3). Due to the models' sheer scale and complexity, the advent of exascale computing offers a tremendous opportunity for scientists to develop a greater understanding of the mantle. To fully utilize the forthcoming resources, Terra Neo is working to design new software with optimal algorithms that permit a scalable implementation towards exascale architectures. To the best of our knowledge, it holds the world record for a linear solver with 10^{13} unknowns.

As the first program of its kind, SPPEXA is unique in numerous ways. It is the first strategic Priority Programme, in that it was initiated by the DFG's Board and did not originate via the standard bottom-up funding methods. SPPEXA also featured many more coordinated activities—such as workshops and doctoral retreats, for example—than is typical of Priority Programmes. It has truly been a multidisciplinary endeavor that involved topics and researchers from informatics, mathematics, and many areas of science and engineering (see Figure 4). Finally, SPPEXA's creation marked the first time that a Priority Programme was synchronized with two other national funding agencies with bi- and trilateral consortia: the Agence Nationale de la Recherche in France and the Japan Science and Technology Agency.

Two culminating events in late 2019 demonstrated the broad impact of SPPEXA's software-related research on all associated fields: (i) an international symposium in Dresden, Germany, at which all consortia presented their results — published in Springer's *Lecture Notes in Computational Science and Engineering*[7] [1], and (ii) a trilateral workshop in Tokyo, Japan, that focused on the convergence of HPC and data science. SPPEXA may have officially come to an end, but the exascale journey continues!

References

[1] Bungartz, H.-J., Reiz, S., Uekermann, B., Neumann, P., & Nagel, W.E. (Eds.) (2020). *Software for Exascale Computing – SPPEXA 2016-2019*. In *Lecture Notes in Computational Science and Engineering* (136). Cham, Switzerland: Springer.

Severin Reiz is a fourth-year Ph.D. candidate under the supervision of Hans-Joachim Bungartz at the Technical University of Munich (TUM), where he has worked as program manager of SPPEXA since December 2018. His research is on high-performance computing (HPC) and machine learning algorithms, in collaboration with George Biros at the University of Texas at Austin. Hans-Joachim Bungartz is a full professor of informatics and mathematics at TUM, dean of TUM's Department of Informatics, TUM Graduate Dean, and a board member of the Leibniz Supercomputing Centre. He is an initiator and coordinator of the SPPEXA program, and his research is on scientific computing with a focus on HPC.

1 http://www.sppexa.de/
2 https://www.top500.org/
3 https://sgpp.sparsegrids.org/
4 https://www.scalasca.org/scalasca/software/extra-p/
5 https://www.humanbrainproject.eu/en/
6 https://terraneo.fau.de/
7 https://link.springer.com/book/10.1007/978-3-030-47956-5

Project	Software	Description
AIMES	SCIL	Scientific compression
ADA-FS	GekkoFS	Data management on file systems
DASH	DASH	Distributed data structures
ESSEX	PHIST GHOST CRAFT RACE ScaMaC	Eigenvalues, sparse matrices, checkpoint/restart, algebraic coloring, matrix collection
ExaDG	deal.II	Partial differential equations w/ adaptive FE
EXA-Dune	DUNE	Distributed numerics
ExaFSA	preCICE Ateles	FSI coupling, discontinuous Galerkin
EXAHD	SG++	Sparse grids
EXAMAG	AREPO	Astrophysics
ExaSolvers	Utopia	Finite elements DSL
EXASTEEL	FE2TI	Multiphase steels
ExaStencils	LFA Lab ExaSlang	Local Fourier analysis, DSL for stencils
ExtraPeak	Extra-P	Performance monitor
FFMK	FFMK	Fault-tolerant microkernels
GROMEX	GROMACS	(Bio)molecular dynamics
MYX	MUST	MPI errors
Terra-Neo	HyTeg waLBerla TerraNeo	Tetrahedral grids, Lattice-Boltzmann, Earth's mantle

Figure 4. *More information about these projects and links to the code repositories are available at http://www.sppexa.de/sppexa-activities/software.html. Image provided by the authors.*

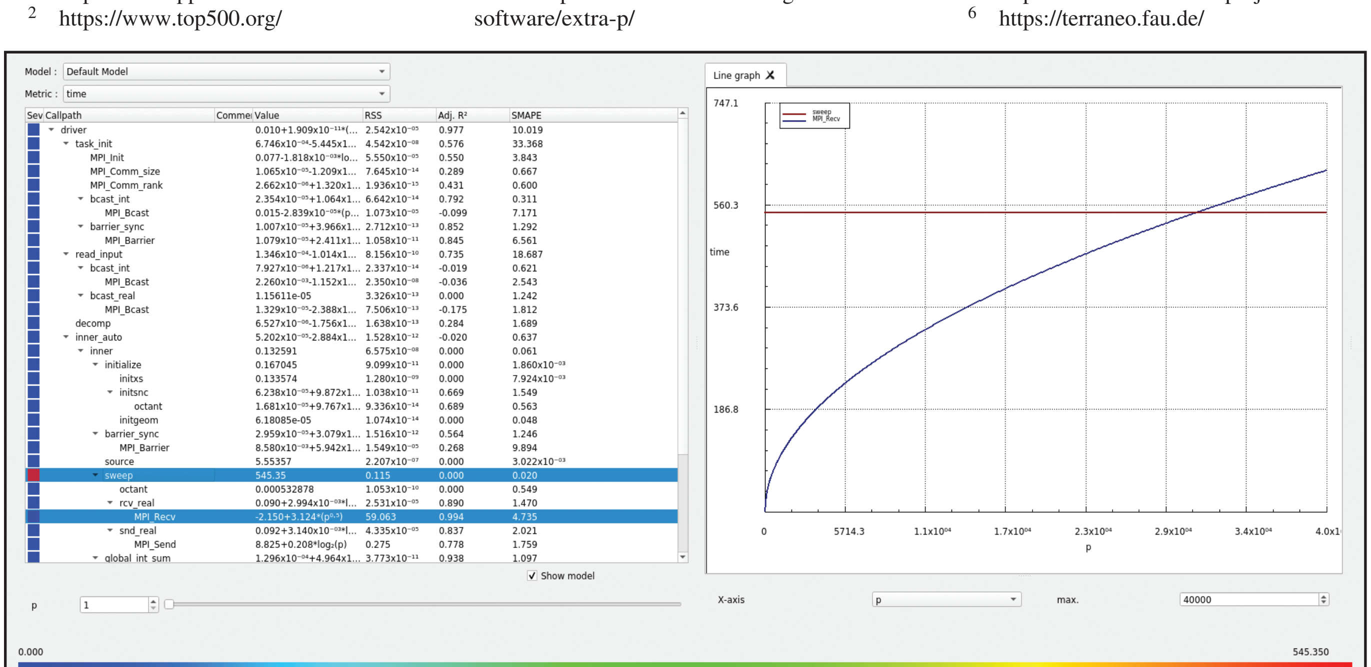


Figure 2. *Performance models generated for call paths in SWEEP3D, a neutron transport simulation. Image courtesy of [1].*

Writing *Learning LaTeX*

By David F. Griffiths
and Desmond J. Higham

The following is a short reflection from the authors of Learning LaTeX, which was first published by SIAM in 1997. The second edition appeared in 2016.

This piece is the second installment of a new SIAM News feature called "From the SIAM Bookshelf," which will periodically spotlight SIAM texts in areas of wide appeal to the greater applied mathematics and computational science communities.

In 1997, we published the first edition of *Learning LaTeX* to provide a useful resource for graduate students, early-career researchers, and other LaTeX beginners. We felt that there was a niche within the prevailing literature for a short and punchy introductory text that covered essential material while avoiding unnecessary detail. The internet was not all-pervading at the time, and the most readily available LaTeX information was present only in comprehensive, expert-oriented textbooks. We aimed for conciseness, accessibility, and concreteness in our book, which displays raw LaTeX code side-by-side with the output.

Producing *Learning LaTeX* was a welcome change from the more restrictive format of mathematical writing. Since we were not competing with the excellent (but somewhat dry) existing LaTeX reference books, we could afford to be highly selective, occasionally incomplete, and lighthearted — all while attempting to avoid misleading our readers and encouraging good style (`\ldots` instead of `....`, `\cos` instead of `cos`, and `\langle..\rangle` instead of `<..>`).

At merely 90 pages long, our inexpensive little text proved to be a popular member of SIAM's portfolio and continued to attract interest well into the internet era, even when multiple online tutorials, beginner guides, symbol lists, and template documents became easily accessible. Many commentators have noted that humans have a penchant for the analogue, as markets still exist for paperbacks, board games, vinyl records, notebooks, and instant cameras. So perhaps a brief hardcopy LaTeX guide retains some value, even when the system generates documents that will only ever live online; for many users, flipping through a short manual in search of an example is likely more satisfying than opening yet another window in their browser. Moreover, because we focused on the absolute basics, the material was less likely to go out of date. For example, *Learning LaTeX* highlighted the difference between "old" LaTeX2.09 and "new, but intermediate" LaTeX 2_ε, which was to be superseded by LaTeX3. But so far, LaTeX3 has yet to emerge. The book's contents thus aged less rapidly than we envisaged.

FROM THE SIAM BOOKSHELF

Over the years, we resisted several attempts by SIAM's book acquisitions team to persuade us to write a new edition. They finally softened us up by highlighting a review in the esteemed *TUGboat* journal (the Communications of the TeX Users Group), which was written 17 years after *Learning LaTeX*'s initial publication. "*I was surprised to find that the book is still in print,*" Boris Veytsman wrote in 2013. "*Thus I ordered a copy for myself and read it.*" He later added that "*it is a testament to (La)TeX and the authors' efforts that this book can still be used as a first LaTeX book or as a basis for a short practical course. It would be also interesting to see a book using the same approach but updated with the new material. Making it as concise and easy to read as this one will be a challenge.*"

On this basis, we signed the contract for a new edition and began preparing an update. Keeping a keen eye on the page count, we added significant new material on packages made available by the American Mathematical Society, including support for typesetting mathematical symbols and multi-line displays. We also included information on the BibTeX program for bibliography generation, the `beamer` class for presentation creation, and the `a0poster` class for poster formation. Figure 1 offers an extract from the material in section 3.5 on equation environments, which provides a sense of the text's content and style.

The second edition of *Learning LaTeX* published in 2016. It clocks in at 103 pages, including a generous index (with the entries `\index{recursion|see{recursion}}` and `\index{Welshman!bearded}`). To maintain the humor-to-content ratio of the first edition, we added an imaginary list of highlights of "LaTeX Through the Years." In the same spirit, we finish here with our predictions for future LaTeX breakthroughs, leading up to the release of LaTeX3 in 2051.

2021: New government legislation on workplace behavior bans `\prod`, `\squeeze`, and `\sin`.

2022: Under current social distancing rules, the second component of every susceptible-infectious-recovered (SIR) model must be typeset as `dI/dt = \beta I \qquad S - \gamma I`.

2023: Ph.D. student assumes that `\boldmath` was designed for theorems with questionable proofs.

2024: An infinite number of monkeys composing LaTeX documents fail to produce a figure in the desired location.

2025: Latest world record for "Smallest Change Necessary to Meet a Journal Page Count Requirement" stands at `\vspace{-0.00001in}`.

2026: Experiments show that the frequency of LaTeX errors is inversely proportional to the time remaining before a grant proposal deadline.

2027: Article disappears in a puff of smoke after mathematician types `\not \exists`.

2028: LaTeX's table-positioning algorithm passes George Marsaglia's *diehard* battery of statistical tests for measuring the quality of a random number generator.

2029: With a single click, Overleaf allows users to submit to a journal, complain about slow refereeing after two weeks, complain about bad refereeing after three months, and automatically resubmit to a lower-ranked journal.

2030: "Least Parsable LaTeX Command" medal awarded to `\righthyphenmin`.

2031: Undergraduates think that `\mho` will convert a statement into what they actually mean.

2032: Campus barber offers a LaTeX special: `\lmoustache`, `\rmoustache`, `\curlywedge`, `\sqcap`, and `\flat\top`.

2033: `\newcommand{BP}{\HUGE +}` turns out to be a big plus.

2034: `\antihistamine` introduced to combat LaTeX allergies.

2035: `\search` hunts down `runaway arguments`.

2036: A 20-page LaTeX document violates several laws of physics when it shrinks to three pages in `twocolumn` format.

2037: 10-year anniversary of the last recorded instance of a journal-ready article that fails to produce an "`Overfull \hbox`" warning.

2038: `Lonely \item` rescued and hand-reared by `\gather`, `\protect`, and `\heartsuit`.

2039: Many LaTeX users agree that `\argh` should be used after making an error.

2040: New LaTeX warning implemented: "`Paragraph ended before you wrote anything of any interest.`"

2041: Plans to reintroduce `\smallskips` back into the wild are scuppered by the predatory instincts of the non-indigenous `\vfill` population.

2042: Tiebreaker question in final of "International Copyeditor of the Year" asks contestants to differentiate between the results of `f: \mathbb{R} \to \mathbb{R}` and the correct, but almost indistinguishable, `f \colon \mathbb{R} \to \mathbb{R}`.

2043: TeXpert demoted to rank of TeXnician after using Microsoft Word in public.

2044: `\fussy` and `\sloppy` file for divorce, citing irreconcilable differences.

2045: `\supset` and `\frown` become increasingly irritated by `\smile`.

2046: In an effort to curb record levels of font inflation in the U.K., Bank of England declares that `\footnotesize` is the new `\normalsize`.

2047: A rogue `\partial /\partial d x` escapes from a computer laboratory, wreaking havoc within the local `x^k` community.

2048: Survey reveals that 87 percent of LaTeX users import lengthy preambles, donated by well-meaning friends, that nobody has completely understood since 1993.

2049: Worldwide mathematical productivity increases dramatically after discovery of a Newton method alternative to the default troubleshooting technique of bisection with `\end{document}`.

2050: BibTeXing of arXiv articles remains a mystery to all.

Enjoy this article? Visit the SIAM bookstore at **https://my.siam.org/Store** *to learn more about Learning LaTeX and browse other SIAM titles.*

David F. Griffiths is a retired Reader in Applied Mathematics at the University of Dundee. He now uses his LaTeX skills to present the results of his genealogy research. Desmond J. Higham is a professor of numerical analysis at the University of Edinburgh. He is a SIAM Fellow and editor-in-chief of SIAM Review.

```
\begin{align}
 y &=  x^4 + 4          \nonumber\\
   &=  (x^2+2)^2-4x^2 \nonumber\\
   &\le (x^2+2)^2    \label{yineq2}
\end{align}
```

$$\begin{aligned} y &= x^4 + 4 \\ &= (x^2+2)^2 - 4x^2 \\ &\le (x^2+2)^2 \qquad (3.3) \end{aligned}$$

The most striking difference is that `align` uses less space around the aligned symbols than `eqnarray`. This environment can deal with more than one column of equations:

```
\begin{align*}
y &=  +a, &  z &=  -a,\\
  &={}+a, &  z &={}-a,\\
  &= 1+b, &  z &= 1-b.
\end{align*}
```

$$\begin{aligned} y &= +a, & z &= -a, \\ &= +a, & z &= -a, \\ &= 1+b, & z &= 1-b. \end{aligned}$$

There is one `&` for each alignment point and one to separate consecutive equations. In general, any odd number of ampersands is allowed.

Figure 1. *Extract from section 3.5 on "Equation Environments," which demonstrates how the general text of Learning LaTeX mixes with the LaTeX examples.*

Summer Schools

Continued from page 5

sessions) to work through some of the calculations. This setup effectively allowed participants to learn the core material in a friendly and interactive environment, which perhaps might not have been so natural at an in-person school.

At the end of the third week, we finalized the research questions—posed by the students themselves—that framed the second half of the school. Small groups of two to five students self-organized around each research question and collaborated with a mentor and/or faculty leader, often calling upon the experts for online consultation and advice. During the program's last days, each group delivered a 25-minute presentation and submitted their reports for the record.

For the sake of brevity, we neglected to detail many additional activities from the summer school's six-week period. Of particular note were the extracurricular sessions on group dynamics, in-depth discussions on diversity and personal success, and fellowship applications. Activities outside of school hours also brought people together, including a movie night on Netflix Party. Further details are available on SIAM News Online.[4]

Conclusions

Is a virtual summer school possible? The answer is clearly "yes." Such an event might even have some significant advantages over an in-person program. For instance, we were able to keep 10 faculty members engaged for the *entire* six-week period (no one-week or cameo appearances). We could also repeatedly involve the experts and maintain their influence over the research directions. Having all students, leaders, mentors, and experts online had a surprising equalizing effect, and the cumulative result of the ensuing interactions led to strong connections. A number of research groups are still meeting to develop their projects, and some social events even occurred in August — including a weekly tea on Wednesday afternoons.

Our motivation for the virtual summer school stemmed from the fact that many graduate students had already arranged internships or other summer research activities that were cancelled due to the COVID-19 pandemic. We wanted to offer these students an intellectual challenge and provide them with a stipend to get through an unprecedented summer. Holding the school virtually was key for this latter point, as we converted the funds for travel and lodging into stipends and accommodated many more students as a result.

We thank the many people who put in an extraordinary effort to make this school a success (see Figure 2, on page 5). It stands as a model for the future.

[4] https://sinews.siam.org/Details-Page/tag/aimmcrn-summer-school

Chris Jones is a research scientist at the Renaissance Computing Institute at the University of North Carolina at Chapel Hill. He received the 2020 SIAM Activity Group on Mathematics of Planet Earth Prize. Hans Kaper, founding chair of the SIAM Activity Group on Mathematics of Planet Earth and editor-in-chief of SIAM News, is affiliate faculty in the Department of Mathematics and Statistics at Georgetown University.

Communication Networks and the Spread of Misinformation

The Misinformation Age: How False Beliefs Spread. *By Cailin O'Connor and James Owen Weatherall. Yale University Press, New Haven, CT, December 2018. 280 pages, $26.00.*

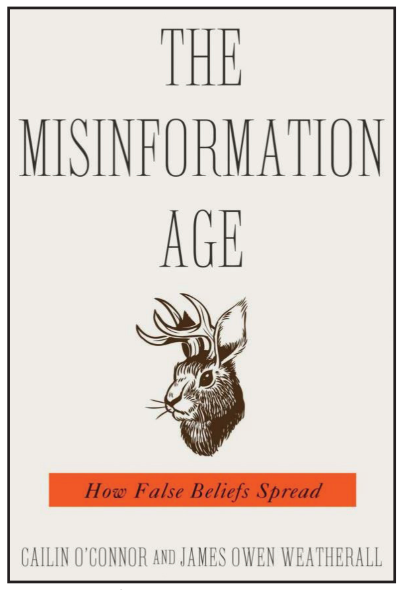
BOOK REVIEW
By James Case

Social scientists Cailin O'Connor and James Owen Weatherall are faculty members at the University of California, Irvine, with secondary appointments at the university's Institute for Mathematical Behavioral Sciences. Their new book, *The Misinformation Age: How False Beliefs Spread*, explores the communication networks through which nonsensical information travels—often with lasting effects—at speeds that depend on the nature of the networks in question.

The introduction and first chapter present a series of anecdotes that concern particular pieces of (mis/dis)information, while the second and third chapters analyze the properties of specific communication networks with the aid of a few simple diagrams. Many of these networks connect scientists, because their connections are comparatively easy to document. Scientific (mis/dis)belief seems to spread in much the same way as other forms of (mis/dis)belief. The fourth and final chapter of *The Misinformation Age* applies middle chapter methods to the role of (mis)information in public life.

The authors' lead anecdote concerns a tale that reached 14th century Europe by way of an English knight named Sir John Mandeville. Upon his return from Asia Minor, he spoke of a tree that bore fruit containing tiny sheep. He claimed to have tasted the flesh of these "vegetable lambs" and found it "wondirfulle." Confirmation soon arrived and Sir Mandeville's story began to circulate among "naturalists," eventually appearing in scholarly books. Not until 1683 did King Charles XI of Sweden direct naturalist Engelbert Kaempfer to undertake the exhaustive search of Asia Minor that was necessary to establish that no such tree existed, or ever had. Belief in "the vegetable lambs of Tartary" overcame the doubts of skeptics to persist among scholars for more than three centuries.

A more recent fake news story appeared in September 2016 on a conservative website known as ETF News (endingthefed.com) under the headline "Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement." While the number of readers who actually believed the story is unknown, it was liked or shared 960,000 times on Facebook between the day it was posted and the election. A skeptic would argue that because the Pope is a public figure, *The New York Times*, *Washington Post*, *Wall Street Journal*, and any number of other media outlets would have reported any such endorsement. However, this story was but one of many false reports. O'Connor and Weatherall calculate that the top 20 fake news stories in the three months before the 2016 U.S. presidential election were liked or shared a total of 8.7 million times on Facebook, while the top 20 genuine news stories during the same period garnered only 7.3 million likes or shares.



The Misinformation Age: How False Beliefs Spread. By Cailin O'Connor and James Owen Weatherall. Courtesy of Yale University Press.

A third anecdote concerns the treatment of stomach ulcers, long believed to be caused by bacteria. This belief went largely unchallenged until 1954, when gastroenterologist E.D. Palmer biopsied the stomachs of more than 1,000 patients without finding evidence of bacteria. The obvious conclusion was that bacteria cannot survive in stomach acid and thus cannot cause ulcers. Instead, physicians believed that ulcers were caused by the acids themselves and could be cured by acid neutralization. In the years that followed, many ulcer patients were "successfully treated" with antacids, though their ulcers displayed a distressing tendency to recur.

Roughly 30 years after Palmer published his results, Australian researcher J. Robin Warren detected a new strain of bacteria in biopsies near the sites of stomach ulcers. His colleague Barry Marshall isolated the new strain, proving that bacteria can in fact dwell in the human stomach. In 2005, the duo received the Nobel Prize in Medicine for convincing their fellow scientists that bacteria can and do cause stomach ulcers in humans.

O'Connor and Weatherall begin their analysis of communication networks by explaining the concept of a model. The models in their book are adaptations of one that economists Venkatesh Bala and Sanjeev Goyal introduced in 1998 [1]. A few years later, Kevin Zollman utilized this model to analyze the networks through which scientists interact [2]. The authors' use mirrors Zollman's work.

Bala-Goyal models describe a collection of "agents" that are attempting to choose one of just two possible conclusions: A or B. They use information gathered by both themselves and others to make their choice. For instance, conclusion A might be that excess stomach acid causes ulcers, which one should treat with antacids; conclusion B might find that ulcers are caused by bacteria and are better treated with antibiotics. Because the conclusions of interest assert the superiority of one action over another, there is no need to distinguish between "conclusion A(resp. B)" and "action A(resp. B)."

Over successive rounds of data gathering, the agents update their tentative conclusions in response to information from the latest round. They initially have little data to analyze and minimal confidence in their verdicts. But as time passes, the agents accumulate, share, and digest data, thus enabling increased confidence in their (still tentative) conclusions.

One can visualize such models as graphs, in which each node represents an agent or group of agents and each edge represents a channel of communication between two nodes. Every node is associated with a number between 0 and 1, called a "credence;" this represents a level of certainty that action B is superior to action A. For example, a node of credence 0.7 indicates that a particular agent is 70 percent certain that B is superior to A, while a credence that is smaller than 0.5 signifies that the agent in question favors the opposite conclusion. With luck, the "credence vector" will converge to a vector of 1s after multiple rounds of data gathering. If not, the process could fail to converge at all or end in a stalemate, with some agents clinging to each belief.

See **Misinformation** *on page 11*

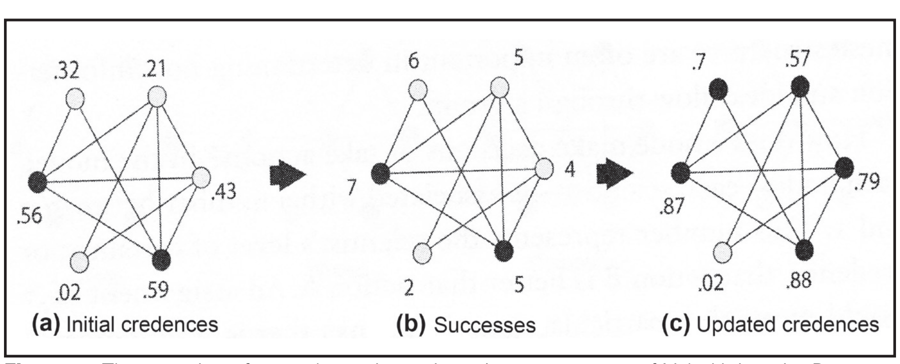


Figure 1. *Three copies of a graph on six vertices. Image courtesy of Yale University Press.*

AN20 Panel Offers Guidance to Future Applied Mathematics Entrepreneurs

By Lina Sorg

Most graduates and early-career professionals in applied mathematics pursue employment in either academia or business, industry, and government. However, many are likely unfamiliar with a third possible career direction: entrepreneurship. While starting a company based on one's strengths and interests may lead to monetary success and personal satisfaction, navigating the complex nuances of the business world can be daunting.

The Second Joint SIAM/CAIMS Annual Meeting,[1] which took place virtually this July, featured an online panel discussion[2] that addressed entrepreneurship within the field of applied mathematics. Speakers Patrick Bangert (Samsung SDS), Frédéric Gardi (LocalSolver), Jeffrey Hoffstein (Brown University and QuantumSafe), and Hector Klie (DeepCast.ai) discussed their individual career paths, spoke candidly about the challenges and gratifications of founding one's own company, and offered advice to future entrepreneurs. Lalitha Venkataramanan (Schlumberger Doll Research) chaired the panel and kept conversation flowing throughout the hour-long session.

CAREERS IN MATHEMATICAL SCIENCES

Bangert, who founded algorithmica technologies[3] in 2005 to bridge the knowledge gap between mathematical applications and industry-based solutions, opened the discussion. He feels that many undergraduate and graduate-level courses implicitly persuade students to become research professors, and warned that numerous academic duties—such as attracting third-party funding, performing administrative tasks, teaching, and grading student work—limit a professor's ability to actually conduct research.

While Hoffstein has enjoyed his stint as a research professor and does have adequate time to work on personal projects, he echoed Bangert's sentiment about academic pressure. "Far too many students out there are indeed told that you must be a researcher," he said. "That is completely untrue." He encouraged students to learn as much pure and applied mathematics as possible, which increases the range of problems on which they can work. Hoffstein noted that much of today's applied mathematics is based on what was completely pure math only 10 years ago. For example, he co-created the NTRU public key cryptosystem in 1996 and established NTRU Cryptosystems Inc. shortly thereafter; the mathematics that inspired his company were completely pure until he and his colleagues found an application.

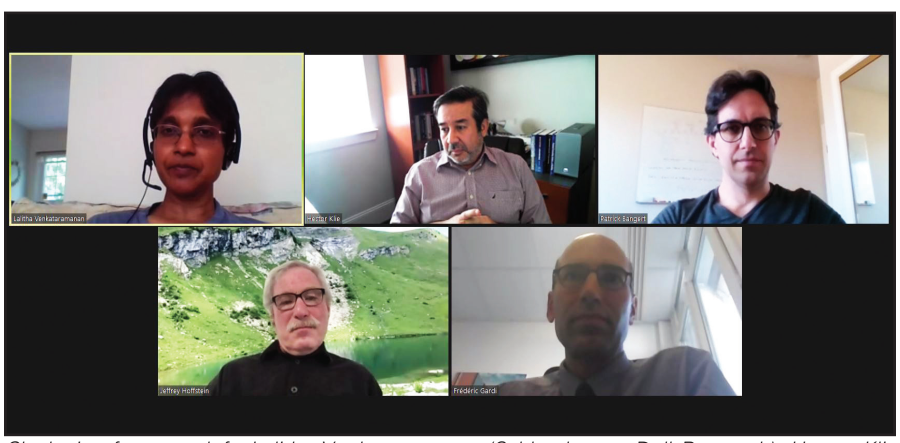
Klie, the chief executive officer (CEO) and co-founder of DeepCast.ai,[4] recommended that all mathematicians who plan to start their own companies first take machine learning courses, master a programming language, and become comfortable utilizing data. Students who complete short-term internships often work with data and are thus at a distinct advantage. "A little experience is a great deal more than none," Bangert said. "What you learn at universities is extremely important and gives you a lot of the background, but the one thing it doesn't teach you is how to apply it in a commercial setting." Internships with commercial businesses also teach students how to report to a boss, succeed in a hierarchical situation, and communicate with colleagues who might not have mathematical backgrounds. All of these skills are necessary when forming and leading one's own company.

Hoffstein urged future entrepreneurs to seek out partners with business sense. "The leadership team is considerably more important than the idea itself," he said. "It's a total disaster to start a company with only technical founders." Furthermore, venture capitalists (VCs) are typically hesitant to invest in a company if none of the founders have business experience. However, one must also be wary of bringing in an outside CEO whose agenda might not directly align with the company's vision.

Nevertheless, a CEO acts as a chief salesperson and no business can exist without paying customers. Bangert indicated that hired CEOs should focus on sales rather than technical matters. While some level of technical capability is useful, it is more important for an effective CEO of a fledgling company to possess a strong communicative ability and extroverted tendencies — traits that help attract VCs. Obtaining VC funding is notoriously difficult, even if a company offers a strong product.

Bangert also provided specific advice for attendees who wish to take on an outward-facing communicative role themselves. "If you yourself are considering doing this, take a few English language and communication classes," he said. "Be able to go out there, speak loudly and confidently, and tell a joke. If you can do some of these things in public, you're on your way to communicating with a nontechnical audience."

See **Entrepreneurs** *on page 12*



Clockwise from top left: Lalitha Venkataramanan (Schlumberger Doll Research), Hector Klie (DeepCast.ai), Patrick Bangert (Samsung SDS), Frédéric Gardi (LocalSolver), and Jeffrey Hoffstein (Brown University and QuantumSafe) partake in a virtual panel discussion about applied mathematical entrepreneurship during the Second Joint SIAM/CAIMS Annual Meeting.

[1] https://www.siam.org/conferences/cm/conference/an20

[2] https://meetings.siam.org/sess/dsp_programsess.cfm?SESSIONCODE=69741

[3] http://www.algorithmica-technologies.com

[4] https://www.deepcast.ai/

Guiding Stress with Cable Networks and the Spider Web Problem

By Guy Bouchitté, Ornella Mattei, Graeme W. Milton, and Pierre Seppecher

When considering a junction of three or more elements that carry a current, we cannot discern in advance how much current will flow through each element; we only know from Kirchhoff's law that the sum of the signed currents will equal zero. This is normally an advantage in an electrical circuit, as the current should flow where it is needed. However, it is disadvantageous when someone is trying to protect a circuit from current overloads or surges. Doing so requires nonlinear elements like current limiters, fuses, and surge protectors. For irrigation, this fact is generally unfavorable because water flow is typically rationed between users; this operation also involves the use of nonlinear elements, such as sluice gates or pipes with limited water-carrying capacity.

The situation is completely different for stress. When four stressed wires meet at a node (or at three nodes, if they are coplanar), the tension in one wire determines the tension in the other three due to the balance of forces. By geometry alone, can one achieve a desired guiding of stress in wire or cable networks under tension? In pentamode metamaterials, exactly four elements meet at the interior nodes; this accounts for their ability to guide stress[1] [3].

Here we focus on finite discrete networks. Structures such as bicycle wheels, suspension bridges, and certain stadiums achieve the desired supporting forces by individually tightening the many spokes or cables and retightening them as needed. In contrast, we aim to create a connected network wherein tightening just one element provides a set of desired supporting forces everywhere. One can achieve this in a simply-connected wire network—where only one (possibly very short) wire is attached to the terminal nodes—by replacing every internal junction where more than four wires meet with a set of junctions, each of which is the meeting point of four or less wires (see Figure 1) [1]. However, applications of this procedure have a significant practical disadvantage: if one wire breaks, the whole network loses its structural integrity, though it may regain its integrity after some finite number of jumps in the points $\mathbf{x}_i$ and changes in the forces $\mathbf{f}_i$ that it supports.

Another important question, which we call the "spider web problem," is as follows. Given a set of n balanced forces $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ at n prescribed points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, when does there exist a wire network under tension (which can be thought of as a spider web, and possibly have many internal junctions) that supports the forces? It suffices to examine webs when no internal nodes exist, which means that one can easily address the query with linear programming [1]. Going one internal junction at a time, we may inductively replace the wires that are connected to this internal junction with a set of wires that pairwise join the nodes that are connected to the one they are replacing (see Figure 2, on page 10). In a sense, this is similar to the "star-delta" transformation in resistor networks, though our result only applies to the stress and not to the elastic response of spring networks where one also monitors the displacement. A more powerful argument indicates that the result holds true even if one starts with a continuum of wires [1]. Thus, if one wants a wire network under tension that is as robust as possible to different loadings, then all of this network's terminals should be connected pairwise. Internal nodes are not needed.

An explicit solution to the spider web problem is available in two dimensions when the points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ comprise the vertices of a convex polygon [6]. This two-dimensional (2D) case corresponds to the points being coplanar—say, in the horizontal plane—with each $\mathbf{f}_i$ pointing along the plane. A web under tension then exists only if the cumulative torque $\Sigma_{k=j}^{m}(\mathbf{x}_k - \mathbf{x}_j) \times \mathbf{f}_k$ in the upward direction is nonnegative for all $\mathbf{x}_j$ and $\mathbf{x}_m$, where the sum is over

See **Spider Web Problem** *on page 10*

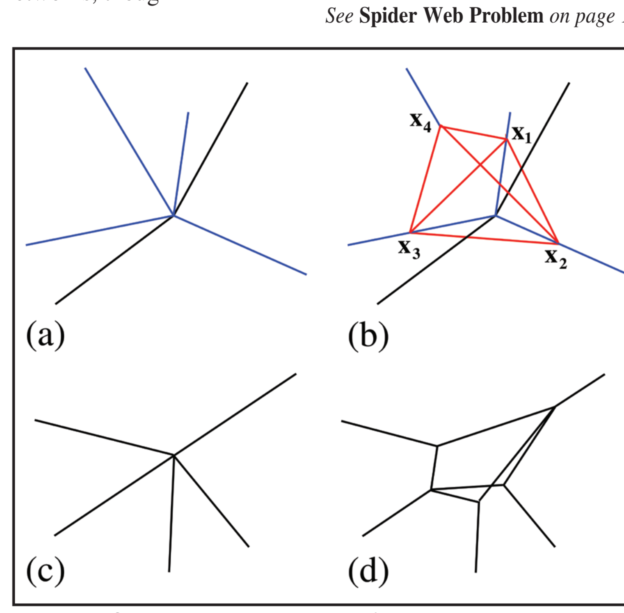


Figure 1. *Steps in the replacement of a many-wire junction under balanced tensions, with a network localized around the junction. No more than four wires meet at any junction, and the network still supports the same tensions in the wires that meet it.* **1a.** *One takes any four wires (in blue) that meet the node—which is presumably at the origin—and are such that they do not all lie on one side of a half-plane.* **1b.** *Next, one picks four points on these wires and superimposes a "tensegrity" network that consists of (i) wires under suitable tension (in red) that join the points pairwise and (ii) four rods under balancing compression that connect these four points and the origin. After superposition, a cancellation of tension occurs in one wire that goes to the origin, which one then removes while maintaining a non-negative total tension in the other wires.* **1c.** *This creates three new junctions where five wires meet.* **1d.** *One can then replace each such junction with a web under tension, which has a maximum of four wires at any of the new junctions. One repeats this procedure until the number of wires that meet at the origin is four or less, and then repeats it at every interior node in the network. Figure courtesy of [1].*

[1] https://sinews.siam.org/Details-Page/pentamode-materials-from-underwater-cloaking-to-cushioned-sneakers

Spider Web Problem

Continued from page 9

all points $\mathbf{x}_k$ going clockwise around the boundary from $\mathbf{x}_j$ to $\mathbf{x}_m$ when viewed from above. One can establish this by considering the Airy stress function $\phi(\mathbf{x})$ in the polygon's interior, denoted Ω. For smooth stresses $\boldsymbol{\sigma}(\mathbf{x})$ with $\nabla\cdot\boldsymbol{\sigma}=0$ in Ω, $\boldsymbol{\sigma}(\mathbf{x})=\mathbf{R}_\perp(\nabla\nabla\phi)\mathbf{R}_\perp^T$. Here, $\mathbf{R}_\perp$ is the 2×2 matrix for a 90-degree rotation with transpose $\mathbf{R}_\perp^T=-\mathbf{R}_\perp$, and $\boldsymbol{\sigma}(\mathbf{x})$ is positive semidefinite in Ω if $\phi(\mathbf{x})$ is convex. For discrete wire networks under tension, this translates to $\phi(\mathbf{x})$ being a convex piecewise linear (piecewise affine) function wherein the wires correspond to its edges and the wires' tension is related to the jump in $\nabla\phi$ across the edge [4]. The convexity of $\phi(\mathbf{x})$ translates to this condition on the torques. One can modify ϕ inside Ω—and correspondingly change the web—as long as the tangent planes of ϕ just outside the polygon's boundary remain unchanged. In particular, researchers can take ϕ as the envelope of these tangent planes and find an equivalent web with no internal loops. They can then extend this result to general 2D webs and replace each convex loop with a locally open web. The area of the loops increases with each replacement until only non-convex loops remain, each of which must have an internal force directed inwards at every vertex where the polygon loop points inward. Thus, if m forces exist in the interior of the convex hull of the points $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$, one can replace any web under tension with a web that has at most m internal loops. A video that demonstrates this process is available online.[2]

This type of web is less likely to be blocked by the edges of solid objects, such as the corner of a wall that may penetrate the convex hull. In the 2D open webs, where the points $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$ are the vertices of a convex polygon, one can move points $\mathbf{x}_i$ backwards and opposite to the direction $\mathbf{f}_i$ while simultaneously maintaining the polygon's convexity and therefore still satisfying the condition for this web to exist. When points $\mathbf{x}_i$ and $\mathbf{x}_j$ merge, one can replace $\mathbf{f}_i$ and $\mathbf{f}_j$ with $\mathbf{f}_i+\mathbf{f}_j$ and continue until just one wire segment supports opposite forces. Retracing the steps provides the open web that supports the original set of forces. Unfortunately, our numerical findings dash the hope that such a procedure may also generate an open web with no closed loops in three dimensions. Initially applying certain forces at a cube's eight vertices and following the same procedure yields forces at the vertices of a distorted convex cube once the game is up. One is thus at a "stuck" configuration, unable to move any of the points $\mathbf{x}_i$ further in the direction of $-\mathbf{f}_i$ while maintaining the "spider web" condition that supports a web's existence [1].

Many open problems remain in this area of research. When can one build webs under tension that avoid given obstacles? If building such webs is possible, is there an associated algorithmic procedure? Can one provide a characterization of "stuck" configurations? Can researchers make progress for finite deformations? That is, given moving points $\mathbf{x}_1(t),\mathbf{x}_2(t),\ldots,\mathbf{x}_n(t)$ and balanced forces $\mathbf{f}_1(t),\mathbf{f}_2(t),\ldots,\mathbf{f}_n(t)$ that are dependent on time t, when does there exist a single web under tension that supports these forces for some given interval of t? Here, single web means a web wherein the angles between wires change with t but the topology and wire lengths do not; all wires also remain under tension and do not collide. Another interesting problem results if some of the wires become slack or collide as t changes. Can one characterize the possible linear responses of spring networks under tension while also considering displacement at the terminal nodes? Researchers have made a complete characterization without the restriction that all springs are under tension [2]. This leads to a complete characterization of the linear elastodynamic response of mass-spring networks while allowing for internal nodes [5].

Acknowledgments: Ornella Mattei and Graeme Milton are grateful to both the National Science Foundation for support through grant DMS-1814854 and the University of Toulon, where this work was initiated.

References

[1] Bouchitté, G., Mattei, O., Milton, G.W., & Seppecher, P. (2019). On the forces that cable webs under tension can support and how to design cable webs to channel stresses. *Proc. Roy. Soc. Lond. Ser. A, 475*, 20180781.

[2] Camar-Eddine, M., & Seppecher, P. (2003). Determination of the closure of the set of elasticity functionals. *Arch. Ration. Mechan. Anal., 170*, 211-245.

[3] Cherkaev, A., Kadic, M., Milton, G.W., & Wegener, M. (2019, June). Pentamode materials: from underwater cloaking to cushioned sneakers. *SIAM News, 52*(5), p. 1

[4] Fraternali, F., & Carpentieri, G. (2014). On the correspondence between 2D force networks and polyhedral stress functions. *Int. J. Space Struct., 29*, 145-159.

[5] Guevara Vasquez, F., Milton, G.W., & Onofrei, D. (2011). Complete characterization and synthesis of the response function of elastodynamic networks. *J. Elastic., 102*, 31-54.

[6] Milton, G.W. (2017). The set of forces that ideal trusses, or wire webs, under tension can support. *Int. J. Solids Struct., 128*, 272-281.

Guy Bouchitté is an emeritus professor of mathematics at the University of Toulon. Ornella Mattei is an assistant professor in the Department of Mathematics at San Francisco State University. Graeme Milton is a distinguished professor of mathematics at the University of Utah. He is a SIAM Fellow and recipient of the SIAM Ralph E. Kleinman Prize, among other distinctions. Pierre Seppecher is a professor of mathematics at the University of Toulon.

[2] https://dx.doi.org/10.6084/m9.figshare.c.4421573

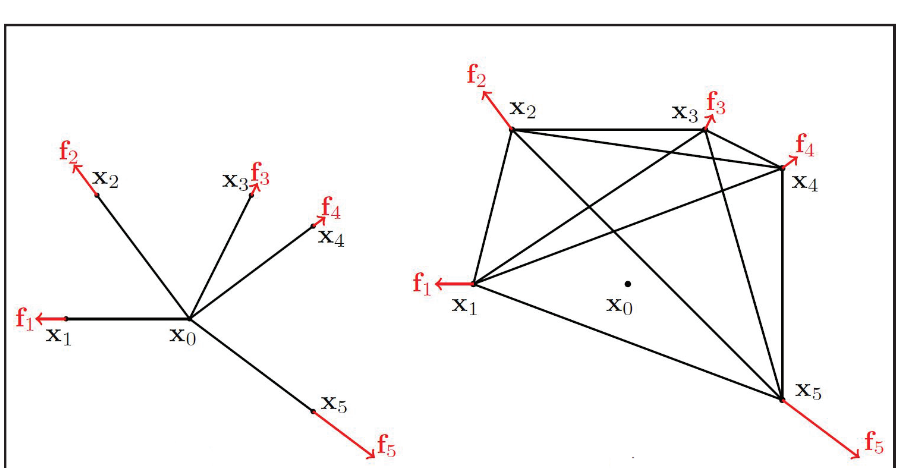


Figure 2. *Removal of one internal node at $\mathbf{x}_0$ that is linked to n other nodes at $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$. If T_{0j} denotes the tension in the wire from $\mathbf{x}_0$ to $\mathbf{x}_j$, then one can remove node $\mathbf{x}_0$ and replace it with a web in which the surrounding nodes are pairwise connected and have tensions $T_{ij}=||\mathbf{x}_i-\mathbf{x}_j||\,c_ic_j(\sum_k c_k)^{-1}$, where $c_j=||\mathbf{x}_i-\mathbf{x}_0||^{-1}\,T_{0j}$. If a pair of these nodes were already connected, one should add this tension to the existing tension. Here, $||\mathbf{x}||$ denotes the length of $\mathbf{x}$. Repeating this procedure allows one to remove all internal nodes. Figure courtesy of [1].*

The Cauchy-Schwarz Inequality via Springs

Here I present a different physical implementation of the idea in [1][1] and [2]; the mathematical portion is exactly the same except for the notations, but I still present it for the sake of self-sufficiency. Let us connect n springs end-to-end, as shown in Figure 1 for $n=3$. Initially, we forcibly hold the connections at some arbitrary positions. Then we release them and let the system settle into the equilibrium configuration. In the process, potential energy decreases:

$$P_{\text{old}} \geq P_{\text{new}}. \qquad (1)$$

The equality holds if and only if the system was already in equilibrium at the outset. I claim that (1) is the Cauchy-Schwarz inequality (in disguise) if the springs are Hookean, i.e., if the tension of the ith spring is in direct proportion to its length: $F_i = k_i L_i$. Indeed, since a Hookean spring's potential energy is $\frac{1}{2}kL^2 = \frac{1}{2}\lambda F^2$—where $\lambda = k^{-1}$ is the spring's "laxness"—(1) amounts to

$$\sum \lambda_i F_i^2 \geq \bar{F}^2 \sum \lambda_i, \qquad (2)$$

where $\bar{F}$ is the common tension of all the springs when the system is in equilibrium. But

$$\bar{F} = \frac{\sum \lambda_i F_i}{\sum \lambda_i}$$

is the weighted average; I omit the verification of this fact, which involves showing that *"laxnesses" add for springs connected in series.*[2] Substituting this expression into (2) gives

MATHEMATICAL CURIOSITIES
By Mark Levi

$$\sum \lambda_i \sum \lambda_i F_i^2 \geq \left(\sum \lambda_i F_i\right)^2.$$

By setting $\lambda_i = x_i^2$ and $\lambda_i F_i^2 = y_i^2$, we get the familiar form of the Cauchy-Schwarz inequality. And if the springs are non-Hookean, with $F = kL^p$ and $p \neq 1$, then (1) amounts to a Holder inequality via an essentially verbatim repetition of the argument in [3].

[1] https://sinews.siam.org/Details-Page/the-cauchy-schwarz-inequality-and-a-paradoxpuzzle

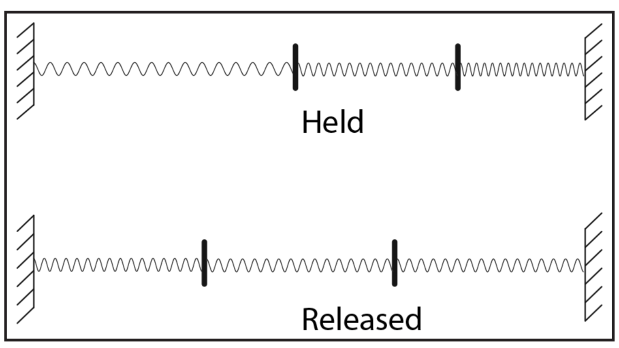


Figure 1. *Potential energy decreases as the system relaxes to the equilibrium. This is the Cauchy-Schwarz inequality for Hookean springs and the Holder's inequality for "polynomial" springs with $F = kL^p$. Figure courtesy of Mark Levi.*

References

[1] Levi, M. (2019, November). The Cauchy-Schwarz inequality and a paradox/puzzle. *SIAM News, 52*(9), p. 11.

[2] Levi, M. (2020). A water-based proof of the Cauchy-Schwartz inequality. *Am. Math. Month., 127*(6), 572.

[3] Levi, M., & Tokieda, T. (2020). A communicating-vessels proof of Holder's inequality. *Am. Math. Month.* In press.

Mark Levi (levi@math.psu.edu) is a professor of mathematics at the Pennsylvania State University.

[2] This is an analog of resistance additivity in electric circuits.

Misinformation

Continued from page 8

Figure 1 (on page 8) displays three copies of a graph on six vertices. The fractions beside the nodes in 1a are the agent's initial credences. Light nodes correspond to agents who plan to take action A because they expect A to outperform B, and dark nodes correspond to agents who plan to take action B. The numbers beside the nodes in 1b indicate the number of times that the actions are successful in a series of 10 independent trials by each agent. Finally, the fractions beside the nodes in 1c represent the updated credences that are obtained by application of Bayes' rule. These credences indicate that all but one agent are fairly certain that B is superior to A after a single round of data gathering. The sole agent in disagreement does not update because he is not in communication with any of the agents that chose alternative B. A second round of testing will likely produce unanimity.

Another of the authors' telling anecdotes concerns Lady Mary Wortley Montagu, whose husband became British ambassador to the Turkish Empire. There she encountered a practice called *variolation* — a primitive version of inoculation that involved scratching a person's arm and rubbing a scab or fluid from a smallpox pustule into the wound. Though a few patients did indeed contract smallpox and die, most experienced a mild form of the illness while developing immunity. A smallpox victim herself, Lady Mary successfully variolated her own young son.

Upon returning to England, Lady Mary sought to popularize variolation among the British aristocracy but met resistance from English doctors. She turned to her friend, Lady Caroline of Ansbach, Princess of Wales, for help. Though Lady Mary's information was accurate all along, the practice did not spread among the English nobility until Lady Caroline's two young daughters were successfully variolated.

O'Connor and Weatherall explain Lady Caroline's influence in terms of a sequence of graphs on seven vertices (agents), arranged in a ring of six around a central "queen bee." The latter communicates with everyone else, but they each communicate with her alone. Should the queen happen to revise an opinion—as she does in Figure 2b—the agents are likely to do so as well. Such "star networks" seldom occur by themselves but often lie hidden within larger networks.

Once identified or created, these subnetworks can be of considerable value to propagandists who seek to influence important events. For instance, the Russian military appears to have made subtle use of the star network concept in their alleged attempts to influence the outcome of the 2016 U.S. presidential election. Facebook has since revealed that Russian-produced political content reached as many as 126 million U.S. users.

A Facebook "group" is meant to facilitate discussion between members, while a "page" is designed for an organization or celebrity to commune with "followers." A "community page" lies somewhere in between. Though the creator attempts to attract followers by posting messages that are of interest to a target audience, registered followers may also post messages that everyone sees.

Well before the 2016 election, the Russians apparently began to create community pages of potential interest to a wide variety of existing affinity groups, including the LGBTQ community, Black Lives Matter activists, gun rights supporters, anti-immigration zealots, and even animal lovers. They did so by posting messages that subtly reaffirmed the target audience's beliefs to gain trust and solidify their position as "queen bee" within a star-like communication network. Only then did they begin to inject a few tenuously related fake news stories into the daily flow of community give and take.

The Misinformation Age treats extremely sensitive material in an entirely scholarly manner, with 27 pages of notes and a 36-page bibliography. It is also highly informative and—at least to this reviewer—a genuine page-turner.

References

[1] Bala, V., & Goyal, S. (1998). Learning from neighbors. *Rev. Econ. Stud., 65*(3), 595-621.

[2] Zollman, K.J.S. (2007). The communication structure of epistemic communities. *Phil. Sci., 75*(5), 574-587.

James Case writes from Baltimore, Maryland.

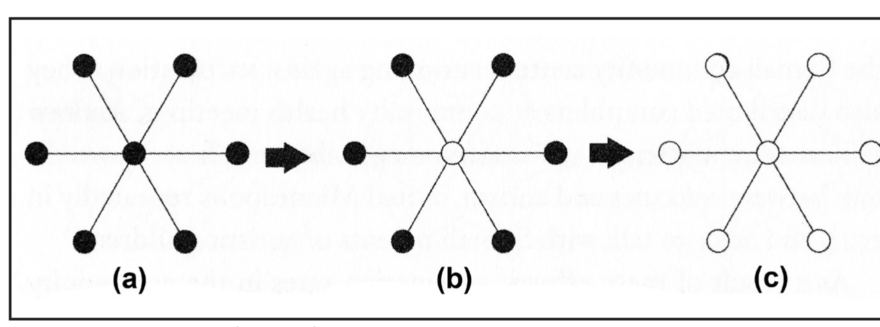


Figure 2. *The spread of a belief in a star network. Dark nodes represent one action and light nodes represent another action. Image courtesy of Yale University Press.*

Professional Opportunities and Announcements

Send copy for classified advertisements and announcements to ***marketing@siam.org****.*
For rates, deadlines, and ad specifications, visit ***www.siam.org/advertising****.*

*Students (and others) in search of information about careers in the mathematical sciences can click on "Careers" at the SIAM website (****www.siam.org****) or proceed directly to* ***www.siam.org/careers****.*

Dartmouth College

Department of Mathematics

The Department of Mathematics at Dartmouth College welcomes applications for a junior tenure-track opening with initial appointment as early as the 2021-2022 academic year. Exceptional cases can merit appointment at a higher rank. The successful applicant will have a research profile with a concentration in applied or computational mathematics. Current research areas in applied mathematics include complex systems, computational social sciences, network analysis, statistical learning, mathematical biology, stochastic processes, uncertainty quantification, partial differential equations, and signal and image processing.

Applicants should apply online at **www.mathjobs.org**, Position ID: **APAM #16169**. Applications received by December 15, 2020 will receive first consideration. For more information about this position, please visit our website: **https://www.math.dartmouth.edu/activities/recruiting**.

Dartmouth is highly committed to fostering a diverse and inclusive population of students, faculty, and staff. We are especially interested in applicants who are able to work effectively with students, faculty, and staff from all backgrounds—including but not limited to racial and ethnic minorities, women, individuals who identify with LGBTQ+ communities, individuals with disabilities, individuals from lower-income backgrounds, and/or first-generation college graduates—and who have a demonstrated ability to contribute to Dartmouth's undergraduate diversity initiatives in STEM research, such as the Women in Science Project, E.E. Just STEM Scholars Program, and Academic Summer Undergraduate Research Experience (ASURE). Applicants should state in their cover letter how their teaching, research, service, and/or life experiences prepare them to advance Dartmouth's commitments to diversity, equity, and inclusion.

The Karen EDGE Fellowship Program

The EDGE Foundation has received an extraordinary gift from Abel Prize winner Karen Uhlenbeck. Her generous gift has been used to establish the Karen EDGE Fellowship Program. Eligible SIAM members are encouraged to apply and spread the word about this unique and exciting opportunity.

Fellowships are available for mid-career mathematicians with full-time employment in the U.S. Applicants must be U.S. citizens or permanent residents with a Ph.D. or equivalent; they must also be members of an underrepresented minority group.

The award consists of $8,000 per year for three years, including funds to support one trip per year to the Institute for Advanced Study in Princeton, NJ (travel only; the institute will provide local expenses) to meet Karen and members of the community. Valid expenses include travel by the Fellow, the Fellow's graduate students, or the Fellow's collaborators for the purpose of advancing the proposed research project; scientific computing; supplies; books; and professional memberships. Teaching buyouts or salary supplements are not permitted. A progress report and financial statement are expected annually within two months of each academic year's end.

The application consists of a personal statement (one page); a research description (two-three pages, not including references); curriculum vitae (two pages); a three-year plan for use of the Fellowship (one page); a budget outline (one page, including travel to Princeton); and current and pending funding support. Applications will be submitted through **https://www.mathprograms.org/db/EDGE/980** and are due by **February 1, 2021**. One awardee will be announced by **May 1, 2021**.

Artificial Intelligence and High-Performance Computing: The Drivers of Tomorrow's Science

By Aparna Chandramowlishwaran

Deep learning (DL)—a specific approach to artificial intelligence (AI) that is based on artificial neural networks—is now recognized as one of the biggest disruptive technologies of the 21st century. After researchers proposed the first neural network with electrical circuits in 1944, the popularity of neural nets rose and waned until recently. Now DL is beginning to revolutionize entire fields and find applications in every conceivable domain of science and engineering.

This transformation was ignited by decades-long advances in high-performance computing (HPC) and the emergence of graphics processing units (GPUs), which make DL feasible in practice. As we move into the next decade, HPC and AI continue to enable and drive each other at an exponential pace. HPC is pushing the limits of AI model complexity—comprised of the number of learned parameters and depth of the mathematical neural network layers used for predictions—for increased accuracy while simultaneously reducing training time. Meanwhile, AI is transforming the way in which we conduct science. The quest for knowledge used to begin with grand theories, and physics-based models still form the core of scientific understanding. But *data-driven surrogate models* are starting to outperform first-principles models in specific tasks, although researchers must address several issues—such as model interpretability and generalization—before such models become commonplace in our science toolbox. Here I summarize recent trends and key challenges that scientists face when using AI to enable scientific breakthroughs.

The most unsettling facet of DL is its *lack of interpretability*. AI surrogates replace highly nonlinear first-principle functions with approximations, which can be orders of magnitude faster than state-of-the-art simulation codes. However, the resulting surrogate model—which has millions or even billions of parameters—is often essentially a "black box," in which case internal representations of the model's learned data (i.e., feature extraction) are unknown. Scientists currently interpret neural networks by observing their external interfaces, not by understanding their internal representations. As a neural net gets progressively deeper and learns increasingly more parameters, it becomes harder for experts to interpret the resulting learned model. This begs the following question: In scientific computing, do researchers really need a billion parameters, especially since it is difficult to reason with these parameters and build confidence models? A promising approach to ameliorate this issue is the incorporation of domain knowledge—such as physical laws, conservation constraints, invariances, and symmetries—in network design and/or training. These domain- or *physics-informed AI models* can potentially require less data and handle more noise while maintaining or improving prediction accuracy by leveraging centuries of advances in scientific knowledge. Although black-box DL models alone are currently not replacements for first-principles models, they already perform remarkably well as preconditioners, software accelerators, and surrogates in complex scientific workflows (see Figure 1).

The core of the DL-based approach to AI is *data*. A unique challenge of scientific DL is both the abundance and lack of readily available data (i.e., ground truth) for training neural nets. In some research areas, instruments commonly generate petabytes of data from one experiment. But in other areas that rely primarily on simulations, generating data to train deep models can be prohibitively expensive. Regardless of how one generates the data, the training process is time-consuming — even on the latest GPU hardware. Although continued advances in computer architecture will accelerate both data generation and training, they are unlikely to keep pace with AI developments. Therefore, researchers must investigate *data-efficient learning* for AI models. Since not all data are equally useful for a given learning task, the data's nature can have a significant impact on the quality of the learned model. One solution is to curate application-specific datasets that can reduce the required data size to train an accurate model. However, while domain-informed AI models and application-specific training datasets can potentially advance model interpretability and reduce complexity, they may also deter generalization. Some key research directions for future exploration include understanding generalization limit tradeoffs versus AI surrogate interpretability, associating confidence intervals with predictions and decision-making to assess model robustness, and quantifying the uncertainty of input data from different scales and disparate sources.

The *portability* of applications and workflows across diverse and increasingly heterogeneous hardware platforms has been a decades-long challenge in scientific computing, and finding the optimal balance between performance and portability remains an active research topic. Driven by commercial AI applications, DL frameworks like TensorFlow,[1] PyTorch,[2] and Keras[3]—which are built upon highly-optimized libraries and software—have risen in popularity and become the "norm" for neural net programming. They are easy to adopt and can abstract underlying evolving AI hardware, parallel algorithms, and architecture-specific optimizations. These frameworks also provide a higher level of programming abstraction, which benefits data-driven scientific computing.

However, this abstraction identifies a new requirement for *AI-integrated scientific workflows*. Existing infrastructure generally does not allow for the seamless integration of the aforementioned DL libraries and frameworks with data from different sources, physical simulations, modeling codes, and *in-situ* analysis tools. Moreover, traditional workflows are based on a human-in-the-loop design that consists of hypothesis construction, repeated experimentation, and exploration of design space. Due to the growing application complexity and hardware diversity, an opportunity exists for *AI-driven scientific workflows*. Future AI might potentially tailor end-to-end application workflows for specialized hardware, with the ability to dynamically adapt as the data and models are refined. This can enable AI to leverage novel solutions that process data close to the source as part of application workflows, thus reducing time to solution.

Data cannot drive the next generation of science if it is inaccessible. The lack of open knowledge bases and curated repositories of data for science applications might arguably be one of the greatest limiting factors in progress when compared to other domains. There is presently no systematic way to contrast the different proposed models for solving the same science problem. An open knowledge base of data and models may both accelerate the pace of model development for scientific applications and allow for *reproducibility* while driving the design of new AI hardware. Applications such as autonomous vehicles, gaming, social networks, and e-commerce are the primary motivators of current AI hardware. Knowledge bases that are tailored for science with exemplar applications could potentially bridge this gap, ultimately allowing science to drive new AI technologies and co-design emerging AI architectures.

Here I describe only some of the core issues and opportunities that researchers may encounter when embracing the DL revolution. DL for science is still in its infancy, and plenty of questions will present themselves in the coming decades. Only time will reveal to what extent DL will augment or replace existing techniques for the solution of real and complex scientific problems.

1 https://www.tensorflow.org/
2 https://pytorch.org/
3 https://keras.io/

References

[1] Obiols-Sales, O., Vishnu, A., Malaya, N., & Chandramowlishwaran, A. (2020). CFDNet: A deep learning-based accelerator for fluid simulations. *ICS '20: Proceedings of the 34th ACM International Conference on Supercomputing*.

Aparna Chandramowlishwaran is an associate professor of electrical engineering and computer science at the University of California, Irvine.

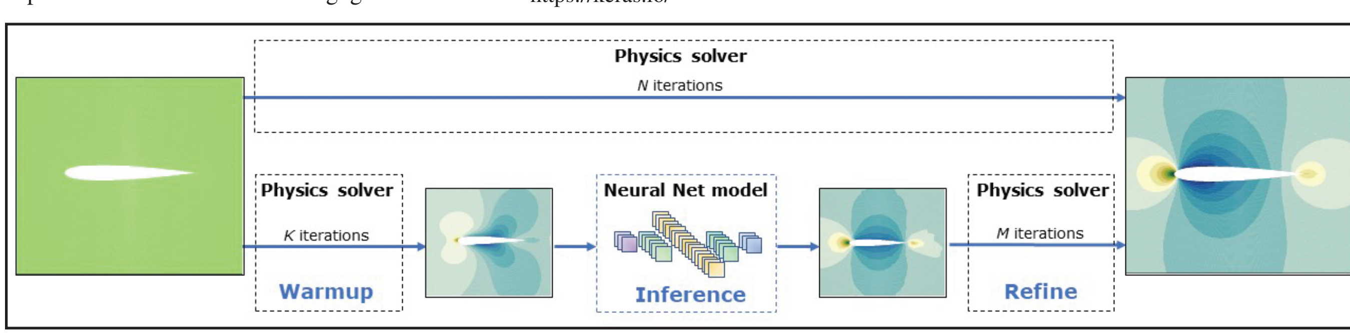


Figure 1. *A sample comparison of the traditional physics solver simulation with a physical simulation and deep learning coupled framework. The latter's objective is to produce the same output as the physics solver (i.e., respect the convergence constraints) while accelerating the convergence process via deep learning. Figure courtesy of [1].*

Entrepreneurs

Continued from page 9

Gardi, who co-founded a global optimization solver called LocalSolver[5] in 2012, filled this role within his company. He admitted that it was difficult in the beginning because he had not studied sales or marketing in school and thus had to learn quickly. "As mathematicians and scientists, we clearly underestimate the ability to sell and communicate," he said. "Ensure that the research and design team is well aligned with business needs to avoid losing time by working on things that don't matter."

Conversation then returned to the role of universities in promoting entrepreneurial skills. "I think universities offer most or all of the necessary components," Bangert said. "Students just need to go out and take them." He reiterated the value of communication courses and suggested that students look for opportunities to meet entrepreneurs who might then become mentors. Hoffstein concurred. "Everything seems to work by links, by person to person," he said. "Absolutely everything is done by personal trust and personal connections." He encouraged students to ask their contacts for introductions to professionals in related specialty areas. Forming a "trust link" with several people often leads to someone with beneficial expertise.

Panelists agreed that amateur entrepreneurs must set financial expectations, decide how much they are willing to sacrifice in personal financial risk, and establish a system of checks and balances to protect their funds from fraud or other unforeseen circumstances. "In order to keep things going, you have to adapt to the times and stay alert," Klie said. "You have to keep the ball rolling." He also recommended that companies start with consulting, with the intent to eventually transition to corporation status. While consulting does earn money and maintain momentum, it is generally not particularly lucrative in the long term. However, it allows companies to explore customers' demands and expectations, experiment with the presentation of their product, and ultimately build a service that consumers actively desire. In addition, Gardi and Klie mentioned that entrepreneurs in this investigatory consulting phase should direct their services towards smaller companies and businesses that do not necessarily have strong internal expertise and are thus more likely to embrace the product in question.

Venkataramanan next steered the conversation to a new topic: issues of intellectual property. In the late 1990s, Hoffstein proactively made a deal with the research office at Brown to ensure that his team's intellectual property was their own and that they had not used any Brown resources in its creation. Nowadays, intellectual property issues are much less archaic and are easily navigable if one is based at a reputable institution or university. The associated research offices might even take care of the patent, work out a fair deal, and help the inventors make contacts.

Bangert expanded upon Hoffstein's explanation, describing how the situation is different in industry. Most companies establish an employment contract, which states that an employee's inventions inherently belong to the inventor's company. Companies are like people in the eyes of the law, in that they can own things and ideas. "There's no conflict in intellectual property in that sense," Bangert said. "It's very clear who owns it, and in no cases is it me." The exception to this rule, of course, is if one's idea is completely unrelated to the employer.

As the discussion wrapped up, Bangert shared some strategies that he uses to keep a calm and stable state of mind. "A large part of my job involves being creative on a short timescale," he said, adding that daily meditation facilitates this process. When Klie finds himself overwhelmed with stress, he partakes in unrelated leisure activities that pull him completely out of a work mindset. Doing so increases his efficiency when he returns to the job.

Upon conclusion of the panel session, attendees had the opportunity to further converse with the experts in online breakout rooms. Despite the challenges of entrepreneurship, panelists urged participants to remain optimistic. All four speakers cycled through various jobs and organizations before arriving at their current positions, reinforcing their stance that there is no concrete path to entrepreneurship. "You can create your own job by becoming an entrepreneur," Bangert said. "Reframe your mindset; it is possible for you, upon finishing graduate school, to start your own company."

Lina Sorg is the managing editor of SIAM News.

5 https://www.localsolver.com/