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

siam news

Volume 49/ Number 8
October 2016

Spatio-Temporal Dynamics of Childhood Infectious Disease

Teasing Out the Limits of Predictability

By Debbie Sniderman

Threats as diverse as HIV, malaria, Zika, Ebola, bioterrorism, and antimicrobial resistance have sparked a growing interest in disease dynamics. Epidemic diseases such as measles have an oscillatory behavior closely analogous to predator-prey dynamics in ecology. Bryan Grenfell (Princeton University) examines the dynamics of measles epidemics in space and time, with special emphasis on how external drivers impact epidemic predictability.

Grenfell attempts to understand predictability and optimize control of infectious diseases; in addition to measles, he studies other childhood epidemics like rotavirus, influenza, respiratory syncytial virus (RSV), and rubella (German measles). These systems are challenging to predict, and prediction is further complicated by the exhibition of strong nonlinear feedback, the evolution of pathogens, and the adaptation of people's behavior to avoid epidemics.

Why Time Matters

Grenfell's approach underlines that simple models can still reveal the essence of epidemic dynamics, despite many biological and social complexities. His time-series TSIR models adapt standard, nonlinear ordinary differential equation (ODE) state-space SIR models to account for measles' operation as a forced oscillator, seasonally driven by the aggregation of children in school. Wavelet spectra help tease out the nonstationary cycles [2]. According to Grenfell, this work is relevant across a surprising range of more complex infections and could be helpful for vaccination and other control measures.

Grenfell advocates the convergence of dynamical systems theory, statistical inference, and detailed epidemic data to explore the diversity of epidemic dynamics. "We should model epidemics from as many pathogens as we can, even non-threatening ones," he says. "We don't know what the next public health challenges will be. Something we don't yet understand may become important."

"There has been a gradual evolution towards synthesizing epidemic dynamics with biology from smaller integrative scales," Grenfell adds. "Increasingly we have to drill down to the molecular and cellular levels, and then back up to individual, community, and population levels. For example, there's currently great progress in predicting seasonal influenza evolution, where all these scales can interact. With so many complexities—network dynamics, complex social networks, individual heterogeneities, within-host dynamics, and battles between viruses and the immune system—the challenge is to find the right level at which to model. Math is the key to closing that gap."

See **Childhood Infectious Disease** *on page 4*

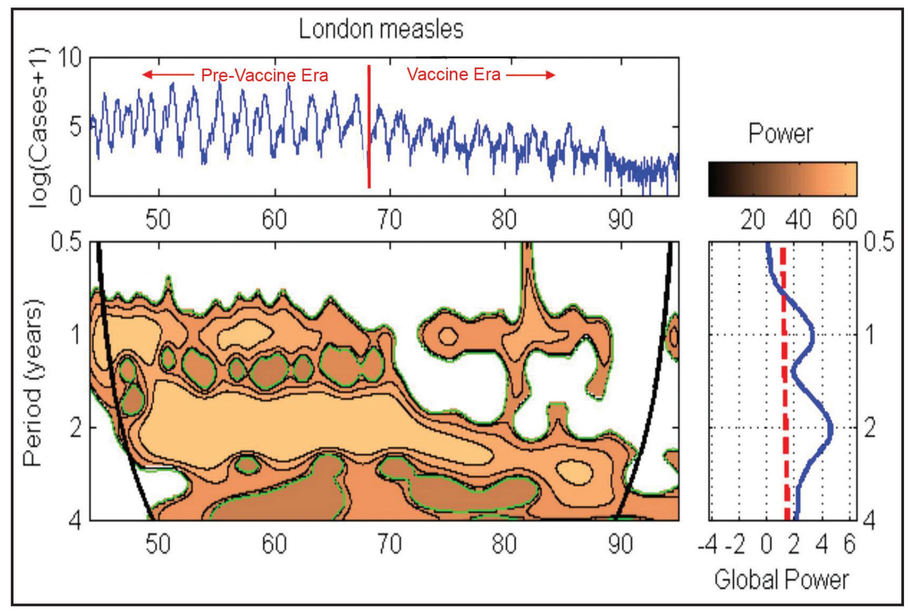


Figure 1. *One and two-year cycles identified by local wavelet power spectra of London measles. Similar non-stationarity effects on cyclicity are seen with vaccination (like reducing birth rate). Image Credit: [2].*

Acoustic Boundary-Condition Dynamics and Internally Coupled Ears

By J. Leo van Hemmen

Humans can discern the direction of a sound source by sensing the time delay between the arrival of sound at each ear. However, in the case of frogs, lizards, and birds, the distance between the ears is too small to make this distinction. Since timely localization of predators and prey is essential to survival, azimuthal sound localization greatly helps animals during both day and night.

For azimuthal sound localization, the time difference between ears is the neuronal cue determined by L/c, where L is the interaural distance and c is the sound velocity. A Tokay gecko, the world's second-largest gecko (lizard), has an L of $\approx$ 2.2 centimeters, which is not very large. Most frogs and birds have even smaller L values. In all cases, merely discerning whether a sound came from the left or the right is not quite what an animal likes doing neuronally. So what is nature to do?

This is where internally coupled ears—ICE for short—come in. More than half of land-living vertebrates are equipped with ICE, which are characterized by a big, air-filled cavity (grey in Figure 1) connecting the left and right eardrums. Let us put the x-axis horizontally through the eardrums in Figure 1 and assume the eardrums are at $x=0$ and $x=L$. For sound localization, the signals at the left and right eardrum have a time difference of $\Delta = L\sin(\theta)/c$, where θ is the sound-source direction, meaning that $\theta = 0$ is directly in front of the animal. The cochlea creates a frequency decomposition and the brain then evaluates Δ; we focus here on Δ, the so-called interaural time difference (ITD), which, due to ICE, is actually *not* what the animal hears.

What knocks on the eardrums is the outside pressure, $p_{\text{out}}(x)$ at $x=0$ and $x=L$, due to the sound source, and it does so (practically) uniformly so that $x=0$ and $x=L$ suffice to specify it. Because of the air-filled interaural cavity between the left and right eardrums in Figure 1, we see coupling mediated by the internal pressure p. The coupling is described [6] by the wave equation $\partial^2 p/\partial t^2 = c^2\Delta_{(3)}p$, where $\Delta_{(3)}$ is the three-dimensional Laplacian equipped with dynamic, time-dependent boundary conditions at the two eardrums that fluctuate under the influence of the external sound source and the internal pressure p.

With a few simplifications, an eardrum can be modeled as a damped linear-elastic membrane with displacement $u(x,r,\varphi;t)$ at either end $x=0,L$ (see Figure 2, on page 3), obeying

$$-\frac{\partial^2 u}{\partial t^2} - 2\alpha\frac{\partial u}{\partial t} + c_M^2\Delta_{(2)}u = [p(x,r,\varphi;t) - p_{\text{out}}(x,r,\varphi;t)]|_{x=0,L}\,. \tag{1}$$

Here α is the damping coefficient, c_M is the membrane's wave-propagation velocity, and $[p(x,r,\varphi;t)-p_{\text{out}}(x,r,\varphi;t)]$ is the total pressure driving the membrane at either end $(\rho_M d \equiv 1)$. In lizards (see Figure 1b), an eardrum can be modeled as a circular disk with Dirichlet boundary conditions $u(a_{\text{tymp}},\varphi)=0$ at the edge $\{r = a_{\text{tymp}}\}$, including a (heavy) cartilaginous sector with boundaries $\varphi = \pm\beta$. It is shown schematically in Figure 2 (on page 3) and transports membrane vibrations to the cochlea.

The setup specified by Figure 2 presents a fascinating and three-fold problem. First, the membrane displacement u in (1) is driven by p_{out} and simultaneously also appears (see Figure 2, on page 3) in p's boundary condition for the three-dimensional wave equation $\partial^2 p/\partial t^2 = c^2\Delta_{(3)}p$ through the famous "no-slip" boundary condition $u = v_x$ with $\rho\partial v_x/\partial t = -\partial p/\partial x$, where ρ is the air's density and v_x is the x-component of the eardrum's velocity. The normal derivative vanishes elsewhere. The only given input is the outside signal p_{out}, and the result is a dynamical system *where boundary conditions for the inside pressure p are part of the system's dynamics*, as in (1). This is the only way to achieve an effective coupling between the left and right eardrums.

See **Internally Coupled Ears** *on page 3*

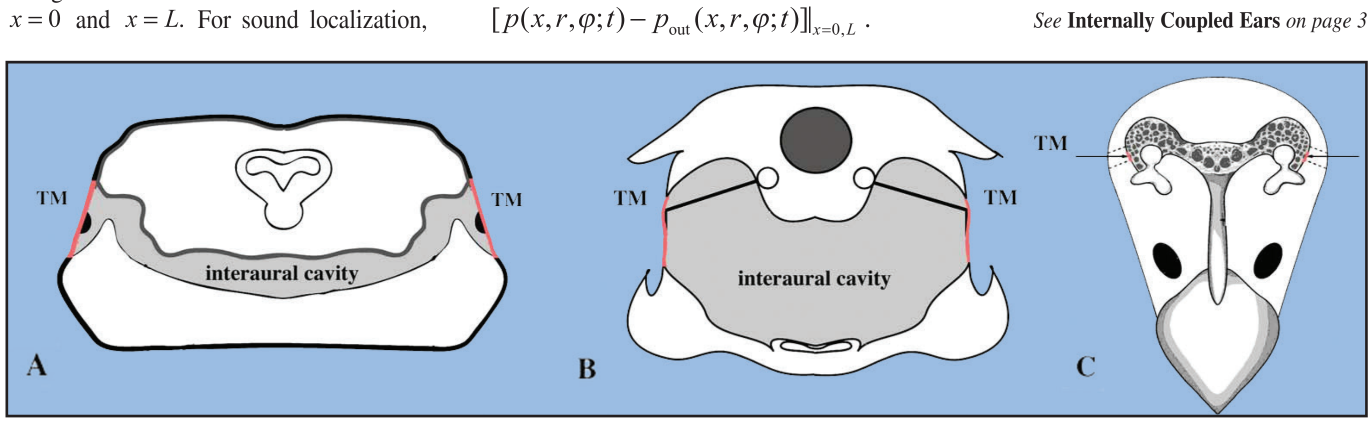


Figure 1. *Internally coupled ears (ICE) in* **(a)** *frog,* **(b)** *lizard, and* **(c)** *bird. Eardrums (TM) are red. Adapted from a figure by Jakob Christensen-Dalsgaard.*

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

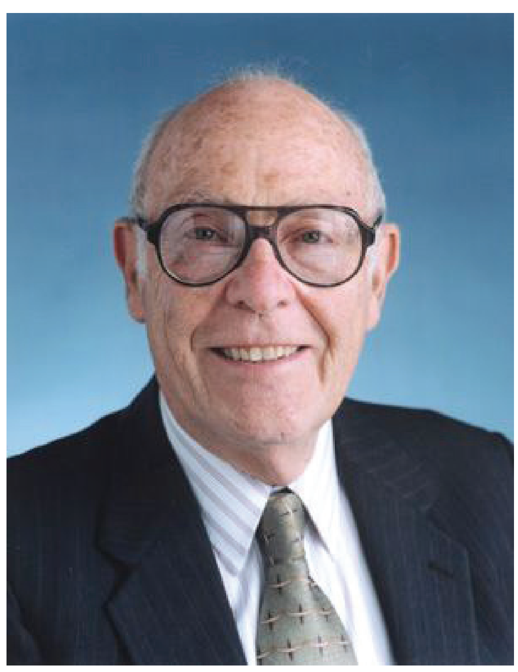
siam.
SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

Obituaries

By Jean Taylor and Carol Handwerker

The mathematics and materials science communities lost a giant when John W. Cahn died of myelodysplastic syndrome (MDS) on March 14, 2016. As tall and imposing as he was physically, his scientific accomplishments made him stand out even more. John is not only *the* Cahn of the Cahn-Hilliard equation and the Allen-Cahn equation; he also developed the theory and some underlying mathematics for the thermodynamics and kinetics of an even broader range of materials phenomena. Spinodal decomposition, critical point wetting, shear coupling, quasicrystals, interface motion, and the importance of symmetries are all indelibly marked by John's hand.

John was born in Germany as Hans Werner Cahn in 1928, though his family soon fled to the Netherlands and then to the United States to escape Hitler. English was the fifth language he learned as a child, though you would never guess it from hearing him speak. John grew up in New York City. He was a proud alumnus of Brooklyn Poly and went on to attend the University of Michigan, although his study there was interrupted by his service in the U.S. Armed Forces in postwar-occupied Japan. John ultimately receive his Ph.D. in physical chemistry from the University of California, Berkeley. He spent his early career in the General Electric (GE) Research Laboratory in Schenectady, NY, a mecca for practitioners in the new field of materials science. When GE turned its interests away from basic research, John became a professor at the Massachusetts Institute of Technology (MIT). His passion for mathematics in the service of materials science has been inspiring to so many of us in the field, through both firsthand experience and his work. After earning her Ph.D., Jean (co-author of this obituary) was an instructor at MIT when she volunteered to give a series of three lectures on soap bubble clusters. John attended them, introduced himself, and a lifetime of collaboration began.



John W. Cahn, 1928-2016. Photo credit: National Institute of Standards and Technology.

John was married to Anne Hessing Cahn, a political scientist with major influence in arms control. When Anne got a job at the U.S. Arms Control and Disarmament Agency (ACDA) in President Jimmy Carter's administration, John—not being particularly happy in the academic milieu at MIT—accepted an appointment as a Senior Scientist at the National Bureau of Standards, which later became the National Institute of Standards and Technology (NBS/NIST). Here, John found the freedom he had enjoyed at GE and flourished, avoiding managerial responsibility like the plague and finding his "students" to be quite teachable. Carol (co-author of this obituary) and her husband John Blendell followed John from MIT to NBS/NIST to work with him, and eventually Carol became John's supervisor. To avoid confusion as to which John was which, Carol usually referred to her husband as "Blendell."

John was truly interested in seeing women succeed in science, and talked about the work of women scientists all the time. Through him, we became acquainted with many female scientists and their work, which enriched our scientific lives. To this day we haven't encountered another scientist who was more positive and encouraging to the women in his field.

Although trained as a physical chemist and transformed into a metallurgist by his mentor Cyril Stanley Smith, John became a member of the National Academy of Sciences in the Applied Mathematical Sciences section at a relatively young age. His awards include a Guggenheim Fellowship at the University of Cambridge from 1960 to 1961, Carnegie Mellon University's Dickson Prize in Science, the Michelson-Morley Prize from Case Western Reserve University, the American Society of Metals (ASM) Albert Sauveur Achievement Award, the Samuel Wesley Stratton Award from the NBS/NIST, the Rockwell Medal, the Harvey Prize from the Israel Institute of Technology, and Gold Medals from Acta Metallurgica, the U.S. Department of Commerce, and the Japan Institute of Metals. John also received the 1998 National Medal of Science presented by President Bill Clinton, the 2001 Emil Heyn Medal from the German Metallurgical Society, the Franklin Institute's 2002 Bower Award, and the 2011 Kyoto Prize in Advanced Technology for his work in materials science. He was a fellow of both ASM and the Minerals, Metals, and Materials Society (TMS), as well as a member of the National Academy of Engineering (NAE) and the American Academy of Arts and Sciences.

Several characteristics contributed to John's scientific success and influence. When someone pointed out an error or omission he had made, he reacted with delight rather than defensiveness. Additionally, John deliberately wrote papers with lots of loose ends. While most people like to write *the* definitive paper on a subject, John preferred writing the first paper defining a field that others would then develop by tying together those loose ends.

After more than twenty years at NIST, John and Anne retired and moved to Seattle to be closer to their three children and many grandchildren. Unfortunately, John's MDS diagnosis came all too soon. His colleagues held a small three-day interfaces conference with John on Bainbridge Island in September 2015 to celebrate how mathematics has permeated materials science, and in November he attended a delightful two-day event at the University of Michigan celebrating Sharon Glotzer's naming as the John W. Cahn Distinguished University Professor of Engineering. After the meeting, he was happy to see that Jean, John Blendell, and Edwin Garcia were pursuing a problem in grain growth that he had reminded us of at the Bainbridge Island meeting. "Aha!" he said. "A posthumous paper!" After having the opportunity to say goodbye to his friends and colleagues, John lived his last days surrounded by his loving family.

Carol A. Handwerker is the Reinhardt Schuhmann Jr. Professor of Materials Engineering and Environmental and Ecological Engineering at Purdue University, performing research in interface motion in multi-component systems, stability of thin films, and sustainable electronics. Jean E. Taylor is professor emerita of mathematics at Rutgers University and a long-term visiting scholar at New York University's Courant Institute. A fellow of the American Academy of Arts and Sciences, she primarily studies equilibrium and growth mechanisms and resulting shapes for crystalline materials.

siam news

ISSN 1557–9573. Copyright 2016, all rights reserved, by the Society for Industrial and Applied Mathematics, SIAM, 3600 Market Street, 6th Floor, Philadelphia, PA 19104–2688; (215) 382–9800; siam@siam.org. To be published ten times in 2016: January/February, March, April, May, June, July/August, September, October, November, and December. The material published herein is not endorsed by SIAM, nor is it intended to reflect SIAM's opinion. The editors reserve the right to select and edit all material submitted for publication.

Advertisers: For display advertising rates and information, contact Kristin O'Neill at marketing@siam.org.
One-year subscription (nonmembers): Electronic-only subscription is free. $73.00 subscription rate worldwide for print copies. SIAM members and subscribers should allow 8 weeks for an address change to be effected. Change of address notice should include old and new addresses with zip codes. Please request address change only if it will last 6 months or more.

Printed in the USA.

siam *is a registered trademark.*

Internally Coupled Ears

Continued from page 1

Secondly, we need to know the complete solution to this linear system of coupled partial differential equations with time-varying boundary conditions, from the start of the sound signal $(t \geq 0)$ onwards. The standard solution, which is practically the only one known, is quasi-stationary and asymptotic $(\alpha > 0)$, in response to a pure tone of angular frequency ω, obtained by splitting off time through the substitution $\exp(i\omega t)$.

Thirdly, it's highly desirable for solutions to be as exact as possible. On the basis of a solid foundation [8], solutions to all three problems have now been found [7]. The solutions are based on a novel time-dependent perturbation theory à la Paul Dirac—in the style of quantum mechanics and in conjunction with Duhamel's principle for the first two problems and the so-called piston approximation—and are largely analytical. They yield a few surprising findings showing why ICE may facilitate sound localization in small animals.

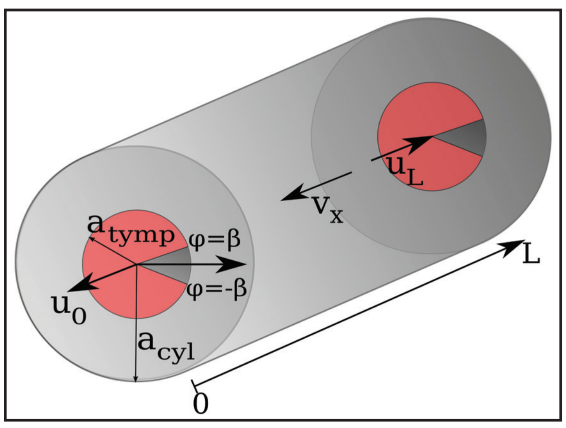


Figure 2. *Effective cylinder representing the interaural cavity. The eardrums are disks with a smaller radius a_{tymp} than a_{cyl} of the cylinder. Grey shaded surfaces are fixed. Image credit: [7].*

Eardrum vibrations and their velocities are amazingly small (< μm and < mm/s for 1 kilohertz (kHz), respectively). An early idea for a simpler problem [1, 2] suggested starting with the original volume $(u \equiv 0)$ and using $\partial p / \partial x = -\rho \partial v_x / \partial t$ as a boundary condition for *p* on both now-fixed tympana. This approach does not, however, account for eardrum fluctuation, and the enclosed volume changes as time proceeds.

It is here that the physics of our problem becomes relevant. Let $\mathbb{1}_{\Delta\mathbb{D}}$ be the indicator function of the interaural-cavity change and provide the restricted Laplacian $\pm \mathbb{1}_{\Delta\mathbb{D}} \Delta_{(3)} \mathbb{1}_{\Delta\mathbb{D}}$ with a plus/minus sign when the eardrum is moving out/inwards. Then we add the time-dependent perturbation $V(t) = \pm \mathbb{1}_{\Delta\mathbb{D}} \Delta_{(3)} \mathbb{1}_{\Delta\mathbb{D}}$ to the unperturbed Laplacian $\Delta_{(3)}$, with $\Delta\mathbb{D}$ as the "excess" volume. Proceed as in quantum mechanics [3], but now start with Duhamel's formula

$$\partial \mathbf{x} / \partial t = [\mathbb{A} + \mathcal{V}(t)]\mathbf{x} \Rightarrow \mathbf{x}(t) = T_t^{\circ}\mathbf{x}(0) + \int_0^t \mathrm{d}t' T_{t-t'}^{\circ}[\mathcal{V}(t')\mathbf{x}(t')] \tag{2}$$

while using the unperturbed semigroup $T_t^{\circ} = \exp(t\mathbb{A})$ operating on a two-component vector $[\mathcal{V}(t')\mathbf{x}(t')]$, where $\mathcal{V}$ contains the perturbation V. The infinitesimal generator $\mathbb{A}$ contains $\Delta_{(3)}^{\circ}$ as the unperturbed and self-adjoint Laplacian with Neumann boundary conditions, i.e., the normal derivative $\partial p / \partial n = 0$ vanishing at the boundary. Start with an "intelligently" chosen zeroth-order state and iterate [4, 7]. By providing the explicit response to a time-dependent (binaural) stimulus, the acoustic boundary-condition dynamics (ABC dynamics) complements Beale's ABC [1, 2].

Air is fairly—though not completely—incompressible; this allows it to carry sound waves. We handle vibrating eardrums approximately, and in so doing exactly, via the piston approximation [7]. We average over the left and right-hand sides of the bases $\mathscr{S} = \{r \leq a_{cyl}\}$ of the cylinder in Figure 2. Once sound arrives, the integral $\int_{\mathscr{S}} \mathrm{d}S\, u(x, r, \varphi; t)\,|_{x=0,L}$ will be nonzero, corresponding to a uniform motion of an effective piston on the left and on the right, the *piston approximation*. The resulting configuration can be treated exactly [7] and errors are relatively small; error estimates can be found elsewhere [4].

Finally, what is the big advantage of ICE in animals? The animals effectively perceive not the interaural difference but the *internal* time and level (intensity) difference, the iTD and iLD, which result from both the external signal *and* internal coupling through the air-filled cavity, as seen in Figures 1 (on page 1) and 2. Here the capital I in ITD or interaural level difference (ILD) refers to interaural; the lowercase i in iTD and iLD refers to internal, and hence to what the animals hear. It turns out that for lower frequencies the fraction iTD/ITD is a flat plateau $\gg 1$, say typically 3-5, so that the interaural distance L effectively becomes much bigger and the localization precision increases as much, at least theoretically. How animals utilize this big profit is subject to hot debate among biologists, who need to interpret the results summarized in [7].

In lizards such as the *gecko*, the effective volume connecting the two eardrums is much bigger than the cylinder with the tympana as endplates. That is, $a_{cyl} \approx 2a_{tymp}$; the piston approximation now makes a lot of sense. The fundamental frequency f_0 is about 1.2 kHz. The lizards' heads are so small that $\mathrm{ILD} \approx 0$ regardless of sound-source direction. Level differences are measured in decibels (dB) as the logarithm of left and right amplitude fraction. Since there is no screening, the logarithm vanishes.

The iTD/ITD plateau occurs for $f < 0.7 f_0$. For higher frequencies, the fraction iTD/ITD then quickly drops to 0 but simultaneously the iLD—as perceived by the animal—can reach, for a direction of $\pm 90^{\circ}$, a level as large as 20 dB (humans can just perceive 1 dB difference.) The maximum of the iLD occurs very near $(< 5\%)$ to f_0 so that it can be used to determine f_0 in a live animal – a notorious problem in the past that can now be solved straightforwardly due to the present theory [7]. Figure 3 provides a summary.

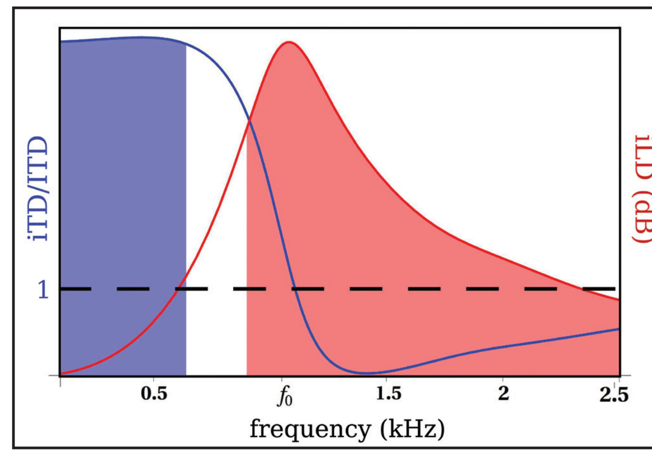


Figure 3. *The eardrum's fundamental frequency f_0 segregates the low and high-frequency domains with different cues, time versus amplitude difference. iLD is a left/right amplitude fraction measured in dB. Image credit: [7].*

What has been said about lizards also holds true for frogs and birds. Both frogs and lizards even have two kinds of cochlear hair cells; one for lower $(< f_0)$ frequencies and sensitive to time differences, the other for higher frequencies $(> f_0)$ and sensitive to amplitude differences. Though there is still hot debate among biologists as to what the animals do with their ICE, the neuronal architecture [5] corresponds well with the biophysics of ICE in having discrete pathways for low and high-frequency hair cells.

Why do mammals refrain from ICE? They may well be small, but they have replaced the air in the air-filled tube with brains. In an evolutionary sense, brains win out over empty space.

References

[1] Beale, J.T., & Rosencrans, S.I. (1974). Acoustic boundary conditions. *Bulletin of the American Mathematical Society, 80*(6), 1276-1278.

[2] Beale, J.T. (1976). Spectral properties of an acoustic boundary condition. *Indiana University Mathematics Journal, 25*(9), 895-917.

[3] Dirac, P.A.M. (1958). *The Principles of Quantum Mechanics* (4th ed.) New York, NY: Oxford University Press.

[4] Heider, D., Vedurmudi, A.P., & van Hemmen, J.L. (2016). *Acoustic boundary-condition dynamics and internally coupled ears*. TUM preprint.

[5] Szpir, M.R., Sento, S., & Ryugo, D.K. (1990). Central projections of cochlear nerve fibers in the Alligator Lizard. *Journal of Comparative Neurology, 295*, 530-547.

[6] Temkin S. (1981). *Elements of Acoustics*. New York, NY: Wiley.

[7] Vedurmudi, A.P., Goulet, J., Christensen-Dalsgaard, J., Young, B.A., Williams, R., & van Hemmen, J.L. (2016). How internally coupled ears generate temporal and amplitude cues for sound localization. *Physical Review Letters, 116*(2), 028101.

[8] Vossen, C., Christensen-Dalsgaard, J., & van Hemmen, J.L. (2010). An analytical model of internally coupled ears. *The Journal of the Acoustical Society of America, 128*(2), 909-918.

J. Leo van Hemmen is an emeritus professor of theoretical biophysics in the Physics Department at the Technical University of Munich (TUM). He also belongs to the Department of Mathematics at TUM, and advocates that a full understanding of the neuronal mechanisms of perception requires comprehension of its underlying biological physics.

Looking Beyond the Black Box in Optimization

By Paul Davis

Optimization problems can be hard for many reasons. But how often is the challenge of such a problem the evaluation of the objective function? How hard can it be to crank out a single number?

Well, maybe not "hard," but certainly time consuming and computationally very expensive. Perhaps the quantity to be minimized can be calculated only by some creaky piece of legacy code. Or its evaluation may demand code packaged in a black box, possibly proprietary. Or it might require using a complex, long-running simulation—or even conducting an entire experiment—for every scenario considered. All of these very real possibilities have arisen as the world's appetite for optimal has grown.

In an invited lecture at the 2016 SIAM Annual Meeting, Stefan Wild (Argonne National Laboratory) addressed this class of optimization problems. He emphasized that the challenge arises because the computational cost of evaluating the objective function overwhelms *every other aspect* of finding an optimum, not because the objective function lacks derivatives.

For example, Wild and Christine Shoemaker (Cornell University) have described the problem of finding optimal pumping strategies to reduce hazardous waste contamination of groundwater: a series of wells, operating over a range of pumping rates, can inject pure water or remove and treat contaminated water. What is the least-cost pumping strategy for 15 wells operating over 30 years, say, to bring the concentration of hazardous agents in the aquifer within the standards of the Environmental Protection Agency? Evaluating a single scenario requires a time-consuming (45+ minutes) groundwater flow simulation.

That simulation code is too complex to permit algorithmic differentiation, which involves processing the code to produce new code able to evaluate specific derivatives of the objective function. Numerical differentiation and genetic algorithms each encounter the same bottleneck of computational cost: the time needed for a single function evaluation—a complete run of the underlying groundwater simulation—far exceeds every other aspect of the computation.

Philosophically, Wild's approach is to "make the most of little," to exploit the values of the objective function at selected points in order to determine a productive direction for the next minimization step. His setting is a model-based trust region, a compact region over which the objective function is well modeled by a quadratic approximation. The trust region grows or shrinks as it moves with the minimization steps.

To "make the most" of the expensive objective function evaluations, Wild chooses the interpolation points so that the approximation error in the quadratic model of the objective obeys Taylor-like bounds. That is, the error in the function approximation is bounded by the square of a measure of the spacing of the interpolation points and the error in approximating the derivative by its first power. Such models are called *fully linear*. Parallel considerations guide the choice of the interpolation basis functions.

Since his presentation also provided quick comparisons to a variety of related approaches, Wild had little time for details that illustrate the efficiencies offered by fully linear models. For example, fitting a unique quadratic in n variables ($n = 15$ pumping rates $+ \ldots$ in the contaminated groundwater example) requires $(n+1)(n+2)/2$ evaluations of the objective. The value of the fully linear model is its reliance on significantly fewer interpolation points while its gradient still points in the direction of a cheaper operating point.

The overhead of fulfilling the geometric conditions on the interpolation points to achieve fully linear models and the effort of building the model of the objective are small change compared to the exorbitant cost of evaluating the objective function.

In pursuit of "lower minimum values with fewer black-box function evaluations," Wild advocates "exploiting as much structure as is known" before employing the algorithmic structure of the fully linear model trust region approach. Least squares

See **Black Box** *on page 6*

Childhood Infectious Disease

Continued from page 1

Fitting Decades of Measles Data with SIR Models

Measles is caused by a highly-transmissible, strongly-immunizing virus that infects individuals for a short time and can cause significant morbidity. An inexpensive vaccine created in 1963 significantly decreased measles incidence throughout the United States and other countries. However, the infection is still a significant killer in many settings, notably parts of sub-Saharan Africa.

SIR models, by definition, divide the population into compartments: susceptible, infected, and recovered/immune. People infected with measles move unidirectionally through the classes. Simple population-level dynamical models of this system include a mass-action term, bilinear in the sizes of susceptible and infected populations with a coefficient that may depend on population size and time. Before vaccination, repeated measles epidemics occurred in large cities, corresponding remarkably closely to the dynamics of a simple oscillator with peaks and troughs. According to Grenfell, SIR models work reasonably well to fit rich measles data sets spanning many years.

Births increase recruitment to the susceptible population. When measles invades a susceptible population, it has a basic reproduction ratio, or R_0, of approximately 18-20 and quickly depletes the susceptible population. An infected individual acquires lifelong immunity after around 10 days of infection.

As the population becomes increasingly immune, more people who do not get measles are indirectly protected. Herd immunity is a key consequence of this nonlinearity – susceptible people are 'indirectly' protected by the population's overall immunity.

Measles can be analyzed at different time scales with wavelet spectra, which allow observation of strong seasonal trends and regular, biennial, or more irregular longer-period epidemics. Wavelet spectra also show dynamic transitions between epidemic periods. In one example, models reveal that when birth rates were high in London, epidemics occurred yearly. As birth rates dropped, a mixture of annual and major biennial patterns emerged (see Figure 1, on page 1).

Complicating Predictability

Chaos, seasonality, and behavioral dynamics complicate modeling. Behavioral complexities include variable willingness to be vaccinated and self-imposed isolation upon becoming sick. Measles in the pre-vaccination era tended to be less complicated by such behavior. This behavior, as well as the relatively high incidence of measles, make pre-vaccination dynamics an especially good testbed for understanding the epidemic clockwork.

In the pre-vaccination period, measles persisted for a long time in large cities because sufficient transmissions in many settings maintained a chain of transmission in interepidemic troughs. Epidemics faded out in populations below the critical community size of 250-300K. In contrast, as vaccinations reduce measles incidence, epidemics are becoming more irregular, with frequent local stochastic extinctions of infection. High levels of immunization above a 'herd immunity' threshold of 90-95% can interrupt measles transmission even in large populations (see Figure 2).

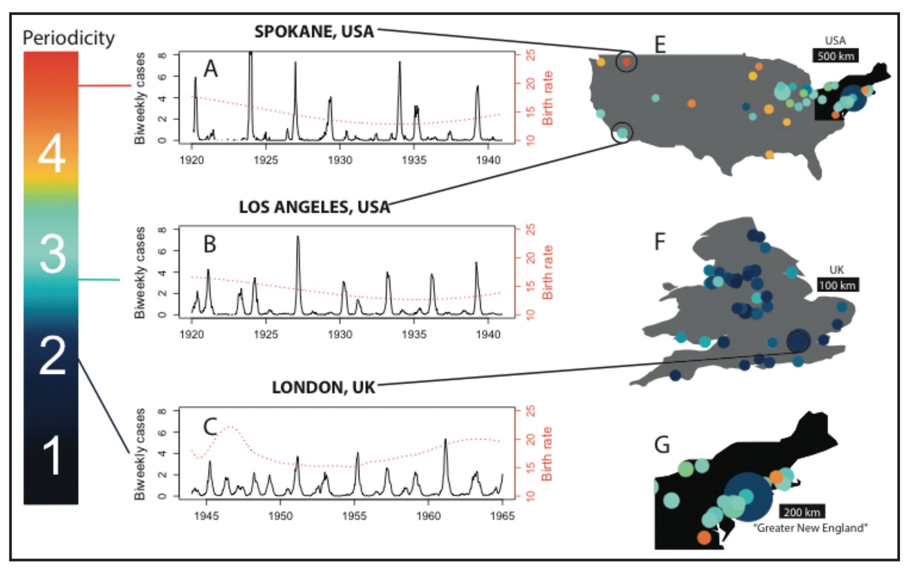


Figure 3. *Epidemic periodicity was longer and more irregular in U.S. cities compared to London. Image Credit: [1].*

Given moderate seasonality of transmission (as in the pre-vaccination United Kingdom data), measles dynamics in large cities corresponded quite well to a relatively predictable limit cycle. Seasonality in the U.K. and the U.S. was largely due to enhanced transmission when children aggregated in school. Typically, outbreaks were high before Christmas and low during the summer holidays.

But nonlinear dynamics can interact with stronger seasonal forcing to create a more complex picture. For example, measles is still a major threat in Niamey, the capital of Niger, where highly seasonal transmission causes much higher amplitude forcing than in the U.K. This has the same effect as powerfully pushing a pendulum, and leads to much more irregular, apparently chaotic dynamics.

Even a subtle increase in forcing can drive dynamic complexities. For instance, summer vacations are slightly longer in the U.S. than in the U.K, and these small increases drove remarkably irregular dynamics during the 1920s and 30s. While London saw biennial and annual cycles, Spokane and Los Angeles experienced three- and four-year cycles, as seen in Figure 3 [1].

"Despite these complexities, we should always try to predict epidemics or understand the limits of predictability where we can't," Grenfell says. "There are great lessons to be learned from predictions made in weather forecasting, which is arguably a more complex system. At the population level, there is evidence for emerging simplicity in epidemiological dynamics. At reasonable scales, the SIR model works well, exposing the limits of predictability. Predict when you can, and understand that problems arise when reaching the limits."

Grenfell believes that in the future a much broader range of expertise will be used to clarify how dynamics of human behaviors, climate change, and pathogen evolution affect epidemic dynamics.

This article is based on an invited lecture by Bryan Grenfell at the SIAM Annual Meeting, which was held in Boston this July.

Acknowledgments: *SIAM News* thanks Tom Kepler (Boston University) for his help with editing this article.

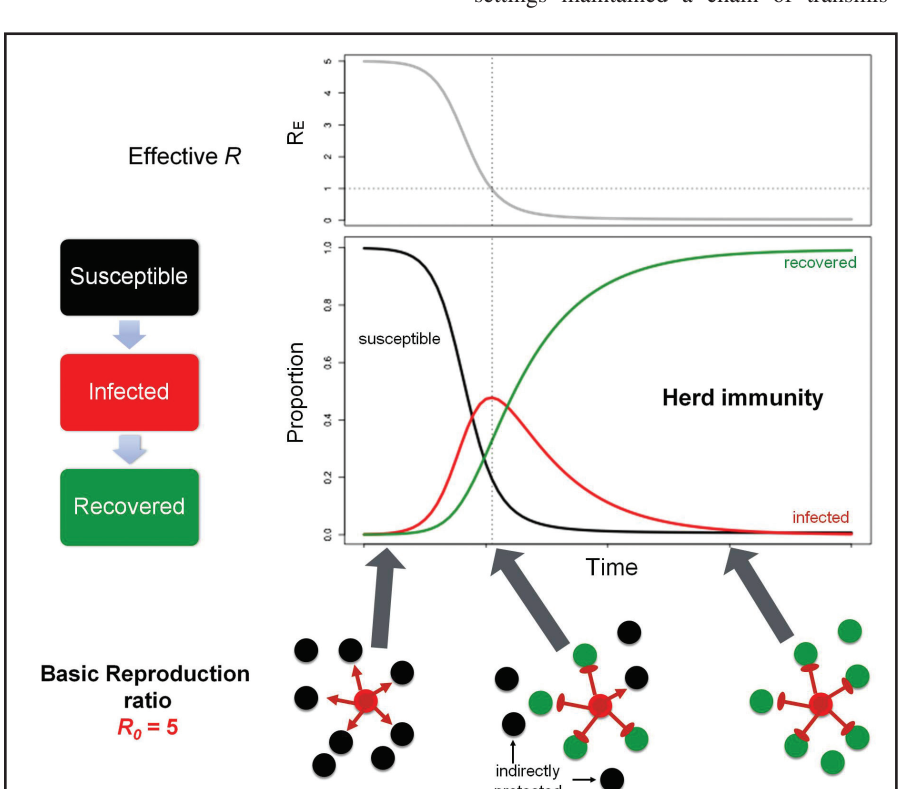


Figure 2. *Herd immunity in the SIR model. Image Credit: Bryan Grenfell.*

References

[1] Dalziel, B.D., Bjørnstad, O.N., Van Panhuis, W.G., Burke, D.S., Metcalf, C.J.E., & Grenfell, B.T. (2016). Persistent Chaos of Measles Epidemics in the Prevaccination United States Caused by a Small Change in Seasonal Transmission Patterns. *PLOS Computational Biology*. http://dx.doi.org/10.1371/journal.pcbi.1004655

[2] Grenfell, B.T., Bjørnstad, O.N., & Kappey, J. (2001). Travelling waves and spatial hierarchies in measles epidemics. *Nature, 414*, 716-723.

Further Reading

Grenfell, B.T., Bjørnstad, O.N., & Finkenstadt, B.F. (2002). Dynamics of Measles Epidemics: Scaling Noise, Determinism, and Predictability with the TSIR Model. *Ecological Monographs, 72*(2), 185-202.

Debbie Sniderman is an applied physics engineer, materials scientist, and CEO of VI Ventures LLC, an engineering consulting company. She can be reached at info@vivllc.com.

Mathematics of Simple Olfactory Search

By Bard Ermentrout

In part 1 of this article, published in the September issue of SIAM News, the author described issues involved with the development of algorithms that animals utilize to locate and follow odor sources and trails. Here, he presents some specific algorithms and explains how one can regard them as interesting dynamical systems.

Bacteria move in response to chemical stimuli, a process known as *chemotaxis*. The molecular mechanisms underlying bacterial chemotaxis are complex, but not well understood. Bacteria shift between directed swimming and "tumbling," a random diffusion-like motion. This is a common strategy among organisms that exhibit odor search; they use *exploitation* (directed motion) when "confident" about odor location and switch to *exploration* (a random search strategy) when they are not. For example, moths pursuing the scent of a mate make big swooping flights (casting) orthogonal to the direction of the wind to locate the odor plume, and then exploit the upwind direction to follow the plume to the source. Animals also exploit other cues, such as rushing water, in addition to cues from the odor itself.

The partial differential equations (PDEs) modeling chemotaxis are well known in the applied mathematics community:

$$\frac{\partial b}{\partial t} = \nabla\Big(D\nabla b - \chi b \nabla C(\mathbf{x},t)\Big), \qquad (\mathbf{x},t) \in R^n \times R^+, \tag{1}$$

where $b(\mathbf{x},t)$ is the density of bacteria, $c(\mathbf{x},t)$ is the concentration of an attractant or repellant, D is the diffusion constant, and χ is the chemotactic coefficient. If $\chi > 0$, then the bacteria will move toward the peak of the concentration.

Since (1) does not account for birth or death, it represents the *density* of bacteria and is proportional to the probability of a given cell being at position x at time t. Thus, (1) is the corresponding Fokker-Planck equation at the single cell level for the following stochastic differential equation:

$$d\mathbf{x} = \chi \nabla c(\mathbf{x},t)dt + \sqrt{2D}d\mathbf{W},$$

where $\mathbf{W}$ is a vector of independent white noise. Thus, the motion of an individual acts as a deterministic strategy (exploitation) that dominates when the gradient of the attractant is large, and as a random walk strategy (exploration) that dominates when the attractant is weak. At the microscopic scale of a single bacterium, the attractant is not a simple smooth function of space. This equation is therefore a very simple idealization.

Imagine a macroscopic organism, like a mouse or a fruit fly, moving continuously in a plane. Its position is (x,y) and its orientation is θ. Its orientation and speed will determine its position:

$$\frac{dx}{dt} = V(x,y,t)\cos\theta$$
$$\frac{dx}{dt} = V(x,y,t)\cos\theta \tag{2}$$
$$\frac{d\theta}{dt} = F(\theta,x,y,t).$$

The functions F,V may only be piecewise-defined and are likely stochastic (hence the explicit time-dependence). The three-dimensional case would involve a third spatial variable, z, and an additional heading angle. The goal of an odor navigation algorithm is to provide details about F and V. In a perfect world, where odor concentration is smooth, the gradient ascent will work quite nicely:

$$\frac{d\mathbf{x}}{dt} = -K\frac{\nabla C(\mathbf{x})}{C(\mathbf{x})},$$

which normalizes the concentration field $C(\mathbf{x})$. Obviously an animal cannot actually compute the gradient, but animals can and do make spatiotemporal comparisons. For example, comparing the odor concentration at spatially-separated sensors (such as the nares (nostrils) of a mouse or the antennae of a lobster) can work nicely in a smooth enough environment. Alternatively, the animal could move its head in a new direction, sniff, and make a comparison to what he sniffed the last time. This is effectively a spatial comparison, just like the two-sensor mechanism. Odor normalization (dividing by the concentration) has a biological correlate: the receptors in the noses of many animals have a logarithmic range of binding constants [1]. Furthermore, the negative feedback circuitry in the olfactory bulb (where the olfactory signals are first processed in mammals) emphasizes differences in concentrations [2]. However, we will not use this normalization in subsequent models as there is a threshold below which odor cannot be detected.

We first consider a simple spatial comparison model where the animal moves with constant speed and only adjusts its heading

See **Olfactory Search** *on page 7*

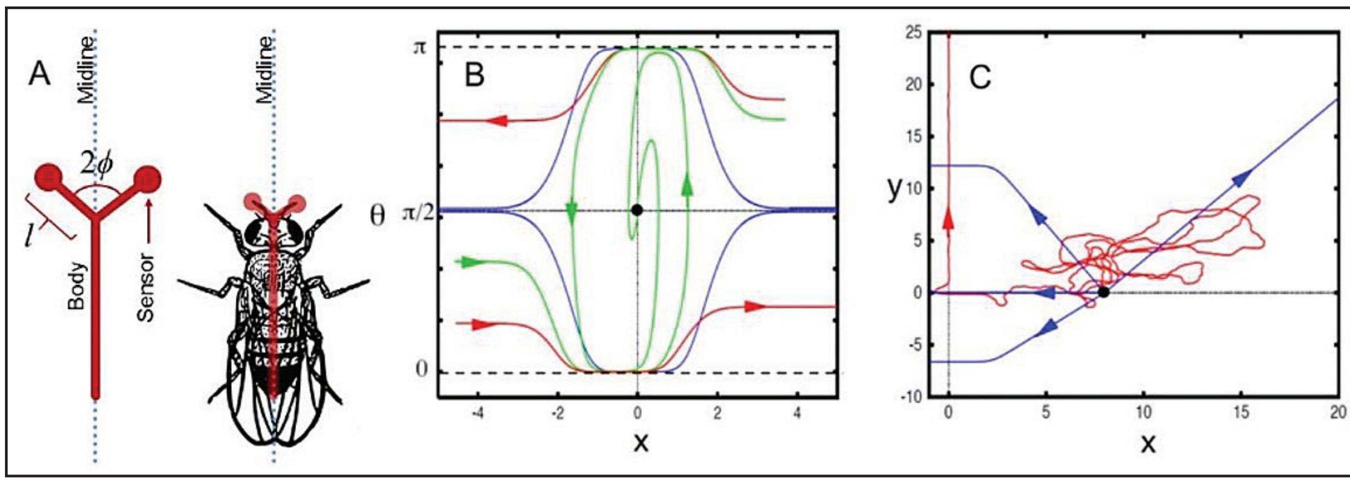


Figure 1a. *Schematic of (3) (on page 7).* **1b.** *Phase plane for this algorithm.* **1c.** *Same algorithm coupled with a correlated random walk. Image credit: James Hengenious.*

ENIAC: The First Electronic Computer's Place in History

ENIAC in Action: Making and Remaking the Modern Computer. *By Thomas Haigh, Mark Priestley, and Crispin Rope. MIT Press, Cambridge, MA, 2016. 360 pages, $38.00.*

BOOK REVIEW
By Ernest Davis

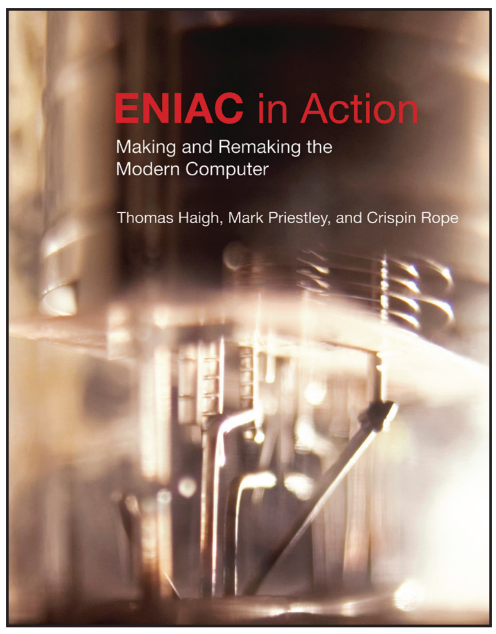
In a blurb on the back cover of *ENIAC in Action: Making and Remaking the Modern Computer*, Paul Ceruzzi (curator of Aerospace Electronics and Computing at the Smithsonian's National Air and Space Museum) claims to have a shelf full of books about the ENIAC, short for Electronic Numerical Integrator and Computer. This seems to be no exaggeration, as there are at least twenty books primarily devoted to the ENIAC, and several dozen more that feature the ENIAC as a major subject. Thus, *ENIAC in Action* is by no means the first word on the subject, and will likely not be the last. However, it is a particularly important, thorough, and balanced account, a major contribution to the history of early computing, and certainly required reading for any student of the subject.

ENIAC in Action looks at its subject primarily from a sociological and technological viewpoint. It is focused on issues such as the ENIAC's intended purpose, the way in which it developed, the tasks for which it ended up being used, and the practical issues involved in its construction. After recounting the history of the ENIAC during its working lifetime, the book proceeds to the ENIAC's afterlife: the protracted, associated patent suit and the contentious debate over its claims to being the first computer. The last chapter is a historiographical discussion of the ENIAC's place in the literature of computing history over the past five decades.

Haigh, Priestley, and Rope have carefully studied and analyzed every scrap of original documentation that they could find. While there are no startling revelations or revolutionary conceptual frameworks in their account, there are lots of fascinating details and some myths put to rest. The struggles of the ENIAC builders to find reliable vacuum tubes are well known; less well known are their problems with more mundane components like resistors and power sources. The cleaning staff was a constant hazard, and eventually the ENIAC operators became practiced in spotting telltale signs that a connection had been knocked out and put back randomly. The idea that the ENIAC, or early computers in general, were used for complex computations on neatly-posed problems with small inputs and outputs turns out to be contrary to the truth. In fact, both the input and the output consisted of enormous decks of punch cards in many problems solved by the ENIAC. Punch cards were also used as external storage to save intermediate states of computation; at times this was a major bottleneck.

ENIAC in Action is striking for the extreme care and thoroughness with which the authors have collected and interpreted historical evidence, and their effort both to avoid letting hindsight drive interpretation and to comprehend how the people involved understood the ENIAC and their relation to it at the time.

Presenting the reader with a clear account of how the ENIAC architecture worked as a computer is not among this book's priorities; the authors undoubtedly—and rightly—feel that this has been sufficiently done elsewhere. The authors are also not concerned with painting character portraits of the people involved. The reader gets a general sense of a collection of very smart people working extremely hard under the immense pressures of wartime and post-war periods. Readers may also perceive the group as contentious, but this could well be the result of many of the protagonists' eventual involvement in a protracted lawsuit.



ENIAC in Action: Making and Remaking the Modern Computer. By Thomas Haigh, Mark Priestley, and Crispin Rope. Image courtesy of MIT Press.

Looming over the entire account is the issue of "firstness." Was the ENIAC "the first electronic computer" or the first electronic computer with certain properties? The question is an unnecessary one, and clearly the authors would have preferred to avoid it, but it drives so much of the subject's literature.

The history of technology has few clear-cut firsts. Many major inventions—the airplane, the telephone, the telegraph—are the subject of competing claims for firstness. The debate regarding the "first computer" is particularly difficult to resolve satisfactorily, for three interrelated reasons.

Firstly, at least in its early years, the ENIAC was not a fixed machine at all in the way we think of machines. One did not run a program on the ENIAC; one brought a problem to the ENIAC team, which used the hardware to procure the answer. W. Barkley Fritz, a member of the team, wrote that the process of designing an ENIAC setup in its original programming method "can be best described as analogous to the design development of a special-purpose computer out of ENIAC component parts for each new application."

Secondly, with most inventions, the inventors were clearly aiming for what we still consider the device's central functionality. The Wright brothers and their competitors were aiming for a heavier-than-air flying machine, Alexander Graham Bell was aiming for a device to transmit sound, and so on. By contrast, the original builders of the ENIAC were not aiming for a "programmable general-purpose computer." They were building a tool for solving particular problems – initially ballistic computations, and later, simulations associated (unbeknownst to them) with the nuclear bomb.

Thirdly, our view of the "general-purpose" computer is very much tied to the now-familiar abstraction from computation theory, with infinite memory. This model is a more or less reasonable idealization of computing technology from the early 1950s onward. But applying it to the machines of the early to mid-1940s is a much more questionable undertaking. Haigh, Priestley, and Rope write the following:

> Discussion of the computational legacies of early computers can easily veer into the counterfactual. . . . This is particularly true when discussion relates to the universality or the Turing-completeness of a machine architecture, since any discussion that begins with the assumption of unlimited time and storage space has already departed irrevocably from the realities of an era in which the overwhelming challenge was to develop reliable and capacious storage. . . . Abstraction is the soul of computer science, but we historians lose something vital if we abstract away from the historical grubbiness of early computer projects, their focus on engineering challenges, and their specific goals and roots in the thinking of the 1940s. For example, Raul Rojas' argument that Konrad Zuse's 1943 Z3 computer was universal was an impressive party trick, but diverged entirely from the way the machine was designed, how it was actually used, or indeed anything that would have made sense in the 1940s.

The clash between the mathematical abstraction and the engineering reality comes to a head in the debate over the significance of John von Neumann's contribution, and in particular of the document he wrote, called "First Draft of a Report on the EDVAC," which laid out the direction for future development. Von Neumann's advocates view this document as the key step in moving from ad hoc calculational hardware to the modern computer, while admitting that von Neumann was not very generous in crediting the ENIAC team. His detractors view "First Draft" as merely a well-written statement of ideas and plans that were already in the minds of ENIAC team members, arguing that the ENIAC would have developed in that direction without von Neumann's intervention.

The question about firstness can be more usefully reframed as, "What was the role of the ENIAC in the development of the electronic computer in the 1940s?" A meaningful answer would involve an examination of all the strands that led to the emergence of the modern computer the following decade. However, a complete answer is not possible; much of the development involved conceptual advances, and tracing the emergence of these concepts would involve reading minds and tracking ephemeral interactions from seventy years ago. Nonetheless, *ENIAC in Action* is a major contribution to understanding the ENIAC's role: the things it accomplished and how it accomplished them, as well as its historical context and resonance.

Ernest Davis is a professor of computer science at New York University's Courant Institute of Mathematical Sciences.

Black Box

Continued from page 4

problems, for example, minimize a sum of squared residuals, some of which may be expensive to compute. The gradient of the objective involves products of individual residuals and their respective gradients. Wild approximates the gradient of each residual term with the gradient of its corresponding quadratic model. For least absolute value problems, Wild notes that because we "know the location of the singularities [we can] use a model-based approximation away from the nuisance points."

And the value of these efficient strategies? "We are closer and closer to being able to optimize virtually everything," Wild says.

Further Reading

Wild, S.M., & Shoemaker, C. (2013). Global Convergence of Radial Basis Function Trust-Region Algorithms for Derivative-Free Optimization. *SIAM Review, 55*(2), 349-371.

Paul Davis is professor emeritus of mathematical sciences at Worcester Polytechnic Institute.

Olfactory Search

Continued from page 5

based on the difference between a left-hand and right-hand sensor:

$$x' = v\cos\theta$$
$$y' = v\sin\theta \qquad (3)$$
$$\theta' = \beta(C_L(x,y) - C_R(x,y)).$$

We assume that the sensors are separated by a distance proportional to l and oriented on either side of the head at angles $\pm\phi$, with respect to the midline (see Figure 1a, on page 5). Thus, $C_{L,R}(x,y) = C[x + l\cos(\theta \pm \phi),$ $y + l\sin(\theta \pm \phi)]$. As a very simple task, we will have the animal follow an infinite trail along the y-axis with a concentration that depends on distance from the trail. For convenience, choose $C(x) = e^{-x^2}$. In this case, y does not enter into the equation for θ and we can study the (x,θ) system via the phase plane:

$$x' = v\cos\theta$$
$$\theta' = \beta\left[e^{-(x+l\cos(\theta+\phi))^2} - e^{-(x+l\cos(\theta-\phi))^2}\right].$$

The only equilibria are $\theta = \pm\pi/2$ and $x = 0$; they are both stable if $\phi \in (0, \pi/2)$. These correspond with following the trail upward and downward respectively. In this case, $dy/dt = \pm v$ and the mouse runs up or down the trail forever. The sets $\theta = 0$ and $\theta = \pi$ are invariant and form the separatrices between the stable equilibria. If an animal gets "stuck" on either of these two lines, it will never reach the stable equilibrium and will move away from the trail forever. More importantly, once x is large enough, $C(x)$ becomes exponentially small, so that θ asymptotically approaches a constant value and the mouse will run off to infinity along that asymptotic heading.

Figure 1b (on page 5) shows several trajectories starting at different initial conditions in the (x,θ) phase-plane; two of them are attracted to the stable equilibrium while others head away, never to return. When the distance is too great, the concentration becomes so small that no corrections in the trajectory are possible. There is a roughly-elliptical region that serves as the basin of attraction to the fixed point. One can understand the approximate shape of this basin by letting l, the sensor distance, get smaller. At the lowest order in l, the resulting equations are integrable and there is an energy contour (shown in Figure 1, on page 5) that provides a sufficient condition for convergence to the trail:

$$v\ln|\sin\theta| + 2\beta\sin\phi e^{-x^2} = 0.$$

Even this simple algorithm invites questions. For example, how well would it perform in a noisy odor environment? How robust is it to sensory noise? How can an animal losing the trail avoid going off to infinity?

As was aforementioned, there is no simple way to model the noisy environment, but we can liken the inputs in the model to what we see in the real world.[1] A simple model for a noisy environment assumes that at each time step, a random event is generated at a rate proportional to the concentration at the sensor position. This position will be either 0 or 1, and the algorithm now becomes driven by the sensor difference $(\pm 1, 0)$. Even the simple left-right differencing algorithm will do quite well if the rate is high enough to ensure that events are not too rare, assuming the initial conditions are near the deterministic basin of attraction. In fact, because of the stochasticity, it is possible to pick up the trail even when the initial data is outside the deterministic basin; the noise causes occasional random turns that may bring the trajectories into the basin. But this advantage works only if the mouse is initially close to the basin. As in the deterministic case, once the distance from the trail is too great, the heading will never change since the events are too rare.

How can we avoid completely losing the trail? As noted above, animals use an "exploration" strategy when they lose the odor or when it becomes too infrequent. Adding a completely random term (not directly related to the concentration) to the deterministic ODE can introduce this type of behavior. For example, we could replace the equation for θ with:

$$d\theta = \beta(C_L - C_R)dt + \sigma dW. \qquad (4)$$

When x is large and the deterministic portion vanishes, θ undergoes a random walk and the variables $x(t), y(t)$ undergo a *correlated random walk* (CRW) since the heading $\theta(t)$ is now a continuous process. The CRW is a search strategy that allows an animal to possibly get back into the basin of attraction of the deterministic dynamics; (4) represents an extremely simple search strategy. One could replace this local exploration with a more general type of search, such as a Levy flight. Figure 1c (on page 5) depicts a correlated random walk, which gets close enough to the trail to be followed. Here we portray the trajectories in the (x,y) plane to clarify the motion. There are four other trajectories without the "noise," starting at the same spatial location but with four different headings. None of them manage to find the trail! The trajectory with the CRW term finds the trail quite readily.

This example considers the simplest spatial comparison model that is continuous in time. A model that incorporates regular casting (moving the head and sampling) yields similar results, except that time is discrete with respect to the sniff cycle. The animal compares the concentration at each sniff to the previous sniff, and then chooses the direction that is towards the greater concentration. This is physiologically different from but mathematically similar to the two-sensor model. Both the two-sensor (or binaral, for animals with noses) mechanism and the alternate sniffing (or casting) model involve odor comparison; in one case comparison is between left and right inputs, while the other case requires some short-term memory. Picking the more salient (in this case, the stronger) stimulus is a task for which the nervous system, and in particular the olfactory system, is well-suited. There is plenty of inhibition (negative feedback) in the early stages of olfactory processing, which can be used to improve contrast between stimuli and even implement a competition in which there is only one winner [2].

In the future, my group wants to determine the underlying neural processes that allow animals to implement various search strategies required to locate food sources, mates, and so on. In summation, locating odor sources in a complex environment is a difficult task with many interesting mathematical features.

References

[1] Hopfield, J.J. (1999). Odor space and olfactory processing: collective algorithms and neural implementation. *Proceedings of the National Academy of Sciences, 96*(22), 12506-12511.

[2] Urban, N.N. (2002). Lateral inhibition in the olfactory bulb and in olfaction. *Physiology & Behavior, 77*(4), 607-612.

Further Reading

Viswanathan, G.M., da Luz, M.G.E., Raposo, E.P., & Stanley, H.E. (2011). *The Physics of Foraging: An Introduction to Random Searches and Biological Encounters.* New York, NY: Cambridge University Press.

Bard Ermentrout is a professor of mathematics at the University of Pittsburgh. He works in many areas of mathematical biology, with a focus on neuroscience.

[1] see "Algorithmically Defining Olfactory Responses in Animals" in the September 2016 issue of *SIAM News*.

A Socially Useful Idea of Privacy

By Paul Davis

In a world awash with data—and plenty of merchants, politicians, and others seeking leverage from that data—we all worry about privacy, even if we lack a sharp definition for it. Perhaps, like pornography, we think we know it when we see it. Or when we lose it.

On the other hand, we give little thought to the validity of analyses based on those vast stores of data. Important policy decisions resting on such data—say, those involving fairness in granting home loans—come and go with little public comment.

In her invited talk at the 2016 SIAM Annual Meeting, held in Boston this July, Cynthia Dwork (Microsoft Research) addressed these and other issues by describing differential privacy, an intuitively appealing but rigorously-defined concept seeing such widespread implementation that its expert practitioners are increasingly in demand. The result of her work and that of her collaborators is a precise formulation of a socially useful idea of privacy.

Massive databases often serve important social ends, such as the identification of links between smoking and cancer, early detection of epidemics from patterns of over-the-counter drug purchases, or extraction of evidence of discrimination from piles of loan applications. Putting aside unauthorized access, the challenge of the last half-century has been preserving the privacy of those whose financial records, in the last example, are part of such a database while permitting analysis of the full body of data for its potential social benefits.

Dwork, a distinguished scientist at Microsoft Research, began her presentation by discrediting some commonly trusted privacy strategies. For example, she stated flatly, “De-identification isn’t.” Anonymity offers no protection if some collective statistics drawn from the data change sufficiently when one individual is removed from the data set. A so-called differencing attack can identify a particular person as a target from one such change, then extract a hitherto unknown fact from another change.

Being a needle in a haystack does not help either. A hacker who holds just a bit of additional genetic information about a few individuals could first determine which of them were represented in a large genetic database, then whether each individual has any of the diseases recorded there. Dwork observed that the availability of “overly accurate estimates of too much data means no individual privacy.”

Some form of privacy protection is essential to entice individuals to risk participating in databases that provide socially useful findings. Data which teaches that smoking causes cancer, for example, could expose individual participants who are revealed as smokers to the risk of higher insurance premiums. How might such protections be formulated, then reliably implemented?

Dwork pointed to an idea advanced in 1977 by Tore Dalenius: privacy-preserving data analysis will not reveal anything about an individual that is not already known. Large-scale analysis of the data remains worthwhile, but such analytic outcomes need to be stable—equally likely—whether or not a given individual is in the population. This sort of stability preserves individual privacy while protecting against overfitting. “Privacy and generalizability are aligned,” Dwork said.

Differential privacy is achieved by applying a mechanism or algorithm to the data before its release; e.g., masking by adding carefully chosen noise. For any two sets of data that differ by a single element, and for all possible outcomes of the privacy mechanism applied to the data in those sets (say, available data entries with noise added), the masking algorithm provides ϵ-differential privacy if the ratio of the probabilities of the noisy data attaining a given outcome on each of the two sets is uniformly bounded above by $exp(\epsilon)$.

Socially, differential privacy offers a probabilistic promise to protect individuals from harm due to their choice to be in a database. Mathematically, the increase of the probability of changing a given outcome by removing one sample from the masked data is bounded; for highly confidential information, that bound can be made close to unity. For example, a health insurer should not be able to detect a change in a count of the smokers in a data set when a prospective customer is removed from it.

This added noise must be chosen with care. For instance, noise that is symmetric about the origin can be averaged away using the answers to repetitions of a fixed query. So-called sensitive queries—those whose answers can vary widely between data sets that differ in only one element—require more noise to conceal differences than do relatively insensitive queries such as counting. A count of smokers, for example, changes by at most one with the removal of a single individual’s data.

Dwork and her colleagues have established that the ubiquitous statistical databases—those with vector-valued data—can provide ϵ-differential privacy when the added noise is symmetric Laplacian with standard deviation proportional to the sensitivity of the potential queries and inversely proportional to ϵ. In Dwork’s expressive phrasing, these databases “smush out the noise” to defend against more sensitive queries or to achieve tighter differential privacy. Differential privacy depends upon the queries and the level of protection, *not* upon the data itself.

One surprising application of these ideas is the reduction of the false discovery rate in adaptive queries of large statistical databases. Applying a differential privacy mechanism prevents queries from revealing anything distinctive about subsets of the underlying database. Since false discoveries are results about a data sample that are not characteristic of the database as a whole, imposing differential privacy prevents their occurrence.

Another application is the provision of reusable hold-out sets for learning. The standard paradigm is to learn (optimize) on a training set, then check against a hold-out set. If the hold-out set is protected by a differential privacy mechanism, then the testing queries reveal nothing about the hold-out set and it can be reused repeatedly.

The algorithmic details and the nuances needed for various important data settings are developing rapidly. Though the devil may be in those details, so are the jobs. Dwork reported that the lack of a competing theory of privacy-preserving data analysis means that those who know differential privacy can find jobs at Apple and the U.S. Census Bureau, among others. Indeed, the chief scientist of the Bureau “is a strong advocate for differential privacy,” a powerful endorsement indeed of the socially valuable, mathematically rigorous discoveries she reported.

Paul Davis is professor emeritus of mathematical sciences at Worcester Polytechnic Institute.

A Bike Wheel and the Gauss-Bonnet Theorem

MATHEMATICAL CURIOSITIES
By Mark Levi

While fixing a punctured bike tire one day, I asked myself whether a wheel, held by the axle and having zero spin, can still turn around the axis. To be specific, let us make the axle describe a closed cone *C*, coming back to its original position and keeping the center of the wheel fixed (the cone need not be circular). Despite the fact that the wheel never spins (the bearings are perfect) it ends up rotated, and through the angle that turns out (apologies for the pun) to equal the solid angle $A(C)$ of the cone *C*.

A geometric theorem about cones can explain this effect. As the wheel's axis traces the cone *C* in Figures 1 and 2, the plane of the wheel remains tangential to another cone C^*, the envelope of the family of planes normal to the generators of *C*. In fact, the plane of the wheel rolls on C^* without sliding, as explained in Figure 2 (this kind of rolling is in complete contrast to the conventional way a wheel rolls on the ground). One can think of the two cones as a bouquet of right angle brackets, as in Figure 1. Speaking loosely, if one cone is sharp, the other is obtuse. The exact relationship between the cones turns out to be the following:

$$A(C)+L(C^*)=2\pi, \qquad (1)$$

where $L(C^*)$ denotes the length of the curve of intersection of the cone C^* with the unit sphere. A physical "proof" of this theorem, along with a rigorous proof, can be found in [1]. The physical "proof"—which led me to discover (1) in the first place—uses imaginary springs and vacuum.

According to (1), $A(C)=2\pi-L(C^*)$; Figure 2 shows that the last expression is precisely the turning angle of the wheel, as claimed.

In fact, the "dual cones" theorem of (1) implies a more general fact, the Gauss-Bonnet theorem – a generalization of the fact that the curvature *k* of a smooth, closed non-self-intersecting planar curve γ satisfies $\int_\gamma k\,ds=2\pi$. According to this theorem (not stated here in its full generality), for a patch *S* of a smooth surface with a smooth boundary γ (see Figure 3), one has

$$\int_S K\,dS+\int_\gamma k\,ds=2\pi, \qquad (2)$$

where *K* is the Gaussian curvature and *k* is the geodesic curvature of γ. According to (2), the bulging of *S* decreases the total geodesic curvature of the boundary, speaking loosely (only the case $K>0$ is discussed here).

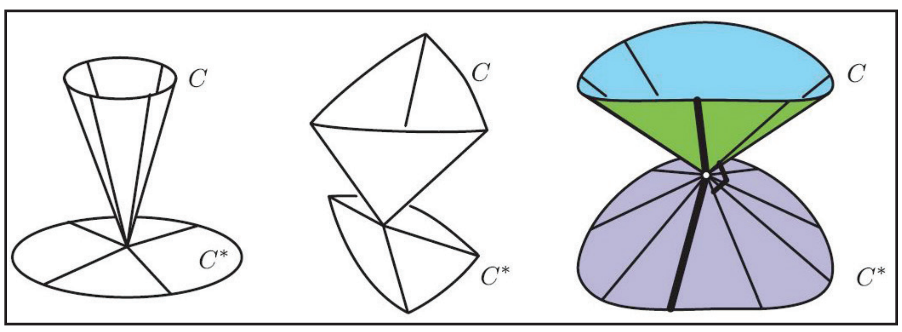


Figure 1. *The dual cone C* is the envelope of the family of normal planes to the generators of C.*

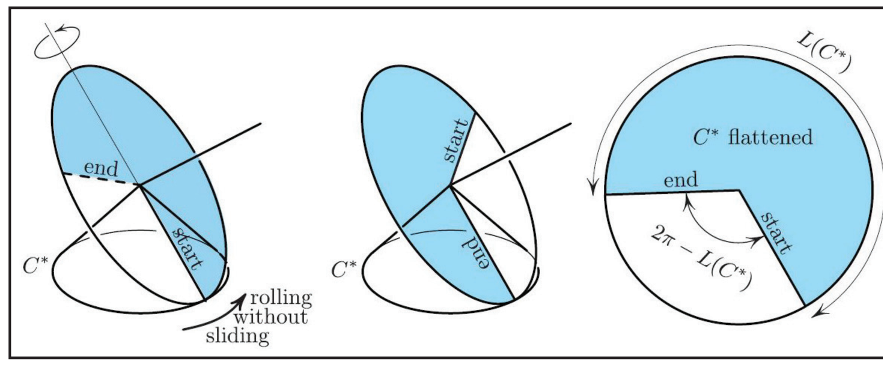


Figure 2. *The plane of the wheel rolls on C* without sliding. Indeed, the wheel is (i) tangential to C* and (ii) the tangency spoke is aligned with the instantaneous angular velocity vector. Hence that spoke is instantaneously at rest, meaning zero sliding. One can think of the disk picking up wet paint from the cone; the angle of the dirty sector on the wheel is L(C), same as the angle of the flattened cone.*

The Gauss-Bonnet theorem boils down to the dual cones theorem (1), as outlined in Figure 3. Starting with the surface patch *S*, we construct the "porcupine" cone *C* of unit normal vectors n (the Gauss cone – Figure 3, right) and consider its dual cone C^*. Intuitively, we can think of walking along γ and carrying a *non-spinning wheel*, keeping its axis (unit vector n) normal to *S*. The wheel's instantaneous angular velocity points along one of the spokes n^* (since the wheel is not spinning on its axis); these vectors transported to a common starting point form the dual cone C^* (see Figure 3, right). Now applying (1) to these cones leads to the Gauss-Bonnet theorem (2).

Indeed, $\int_S K\,dS=A(C)$ by the definition of *K*, and it only remains to explain why $L(C^*)=\int_\gamma k\,ds$. Omitting the details, I will only mention that $L(C^*)$ measures the angle by which n^* rotates in the tangent plane sliding along γ, from the point of view of the observer walking around γ. But n^* and $\dot\gamma$ (the tangent to γ) rotate by the same amount, since the angle between them is unchanged after one traversal of γ; the latter angle is $\int_\gamma k\,ds$ (by the definition of *k*), which explains why $\int_\gamma k\,ds=L(C^*)$ and why (2) indeed reduces to (1).

The wheel provides us with a mechanical interpretation of parallel transport (each spoke undergoes parallel transport if the wheel does not spin on its axis). Further

See **Gauss-Bonnet Theorem** *on page 11*

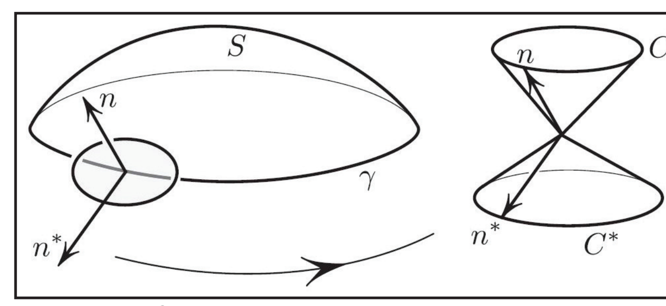


Figure 3. *The Gauss-Bonnet theorem and its connection to (1).*

Professional Opportunities and Announcements

Send copy for classified advertisements and announcements to: ***marketing@siam.org****; For details visit* ***www.siam.org/advertising****.*

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

Institute for Advanced Study

School of Mathematics

The School of Mathematics at the Institute for Advanced Study has a limited number of memberships with financial support for research during the 2017-18 academic year.

The School frequently sponsors special programs. However, these programs comprise no more than one-third of the memberships so that each year a wide range of mathematics is supported.

Candidates must give evidence of ability in research comparable at least with that expected for the Ph.D. degree, but otherwise can be at any career stage. Successful candidates will be free to devote themselves full time to research.

About half of our members will be postdoctoral researchers within five years of their Ph.D. We expect to offer some two-year postdoctoral positions.

Up to eight von Neumann Fellowships will be available for each academic year. To be eligible for the von Neumann Fellowships, applicants should be at least five, but no more than 15, years following the receipt of their Ph.D.

The **Veblen Research Instructorship** is a three-year position in partnership with the Department of Mathematics at Princeton University. Three-year instructorships will be offered each year to candidates in pure and applied mathematics who have received their Ph.D. within the last three years. Usually the first and third year of the instructorship will be spent at Princeton University and will carry regular teaching responsibilities. The second year is spent at the Institute and dedicated to independent research of the instructor's choice. Candidates interested in a Veblen Instructorship position may apply directly at the IAS website (**https://application.ias.edu**) or they may apply through MathJobs. If they apply at MathJobs, they **must** also complete the application form at **https://applications.ias.edu**, but do not need to submit a second set of reference letters. Questions about the application procedure should be addressed to **applications@math.ias.edu**.

Also, the School of Mathematics is looking for highly-qualified applicants in the field of computer-assisted formalization of mathematics, univalent foundations, and homotopy type theory and is expecting to offer two or more memberships in this area.

In addition, there are also **two-year postdoctoral positions** in computer science and discrete mathematics offered jointly with the following institutions: The Department of Computer Science at Princeton University (**http://www.cs.princeton.edu**), DIMACS at Rutgers, The State University of New Jersey (**http://www.dimacs.rutgers.edu**), and the Simons Foundation Collaboration on Algorithms and Geometry (**https://www.simonsfoundation.org/mathematics-and-physical-science/algorithms-and-geometry-collaboration/**).

School term dates for 2017-18 academic year are: term I, Monday, September 25 to Friday December 22, 2017; term II, Monday, January 15, 2018, to Friday, April 13, 2018.

During the 2017-18 year, the School will have a special program on Locally Symmetric Spaces: Analytical and Topological Aspects. Akshay Venkatesh of Stanford University will be the Distinguished Visiting Professor.

The topology of locally symmetric spaces interacts richly with number theory via the theory of automorphic forms (Langlands program). Many new phenomena seem to appear in the non-Hermitian case (e.g., torsion cohomology classes, relations with mixed motives and algebraic K-theory, derived nature of deformation rings). One focus of the program will be to try to better understand some of these phenomena.

Much of our understanding of this topology comes through analysis ("Hodge" theory). Indeed, harmonic analysis on locally symmetric spaces plays a foundational role in the theory of automorphic forms and is of increasing importance in analytic number theory. A great success of such harmonic analysis is the Arthur-Selberg trace formula; on the other hand, the analytic aspects of the trace formula are not fully developed, and variants such as the relative trace formula are not as well understood. Thus, analysis on such spaces, interpreted broadly, will be another focus of the program.

University of California, Davis

Department of Mathematics

The Department of Mathematics at the University of California, Davis invites applications for an Assistant Professor (tenure-track) faculty position in the areas of mathematical optimization and data analytics starting July 1, 2017.

Minimum qualifications for the position include a Ph.D. degree or its equivalent in the mathematical sciences or a related field and excellent potential for performance in teaching and research. Duties include mathematical research, undergraduate and graduate teaching, and departmental, university, and professional service. Candidates are expected to engage in interdisciplinary research within the UC Davis Data Science Initiative.

Additional information about the department may be found at https://www.math.ucdavis.edu/.

Applications will be accepted until the position is filled. For full consideration, completed applications should be received by December 15, 2016. To apply: submit the AMS Cover Sheet and supporting documentation electronically through http://www.mathjobs.org/.

The University of California, Davis is an Affirmative Action/Equal Opportunity Employer with a strong institutional commitment to the achievement of diversity among its faculty and staff.

California Institute of Technology

Department of Computing and Mathematical Sciences

The Department of Computing and Mathematical Sciences (CMS) at the California Institute of Technology invites applications for the position of lecturer in computing and mathematical sciences. This is a (non-tenure-track) career teaching position, with full-time teaching responsibilities. The start date for the position is September 1, 2017, and the initial term of appointment can be up to three years.

The lecturer will teach introductory computer science courses including data structures, algorithms, and software engineering, and will work closely with the CMS faculty on instructional matters. The ability to teach intermediate-level undergraduate courses in areas such as software engineering, computing systems, or compilers is desired. The lecturer may also assist in other aspects of the undergraduate program, including curriculum development, academic advising, and monitoring research projects. The lecturer must have a track record of excellence in teaching computer science to undergraduates. In addition, the lecturer will have opportunities to participate in research projects in the department. An advanced degree in computer science or related field is desired but not required.

Please view the application instructions and apply online at https://applications.caltech.edu/job/cmslect.

The California Institute of Technology is an Equal Opportunity/Affirmative Action Employer. Women, minorities, veterans, and disabled persons are encouraged to apply.

University of North Carolina at Greensboro

Department of Mathematics and Statistics

The University of North Carolina at Greensboro (UNCG), Department of Mathematics and Statistics seeks applications for one tenured/tenure-track position in numerical partial differential equations at the Associate Professor rank beginning August 1, 2017. Candidates must hold a Ph.D. in mathematics or a closely-related discipline and have qualifications to be hired at the Associate Professor rank. Tenure may be offered at the associate professor rank depending on the selected candidate's qualifications. Competitive applicants will have research expertise that strengthens our Ph.D. program in computational mathematics and be ready to direct Ph.D. students. Successful applicants will be expected to excel in teaching, maintain a vigorous research program, seek external research funding, contribute to the interdisciplinary mission at UNCG, and educate a diverse group of undergraduate and graduate students from various backgrounds. Application materials should be submitted electronically to https://www.mathjobs.org/jobs/UNCG. Review of applications will begin on December 15, 2016, and will continue until the position is filled.

UNCG is especially proud of the diversity of its student body, which is 43% ethnic minority (http://admissions.uncg.edu/discover-about.php). UNCG has been designated as a Minority-Serving Institution by the U.S. Department of Education. We seek to attract a diverse applicant pool for this position, especially women and

See **Professional Opportunities** *on page 11*

Gauss-Bonnet Theorem

Continued from page 9

details on this can be found in [1], and the more standard treatments of the Gauss-Bonnet theorem in [2] and [3].

All figures are provided by the author.

Acknowledgments: The work from which these columns are drawn is partially supported by NSF grant DMS-9704554.

References

[1] Levi, M. (1994). A "bicycle wheel" proof of the Gauss-Bonnet theorem, dual cones and some mechanical manifestations of the Berry phase. *Expo. Math, 12*, 145-164.

[2] Pogorelov, A.V. (1959). *Differential Geometry*. (P. Noordhoff & N.V. Groningenm, Trans.) (pp. 160-161).

[3] Spivak, M. (1999). Chapter 6 in *A Comprehensive Introduction to Differential Geometry*. (Volume 3). (3rd ed.). Houston, TX: Publish or Perish, Inc., Houston.

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

samsi

NSF·Duke·NCSU·UNC

Postdoctoral Fellowships for 2017

The Statistical and Applied Mathematical Sciences Institute (SAMSI) is soliciting applications from statistical and mathematical scientists for up to **6** postdoctoral positions for the SAMSI Research Programs for 2017-2018: **Program on Mathematical and Statistical Methods for Climate and the Earth System (CLIM)** and **Program on Quasi-Monte Carlo and High-Dimensional Sampling Methods for Applied Mathematics (QMC)**. Appointments will begin in August 2017 and will typically be for two years, although they can also be arranged for one year. Appointments are made jointly between SAMSI and one of its partner universities, where teaching opportunities may be available. The positions offer extremely competitive salaries, travel stipend, and health insurance benefits.

Criteria for selection of SAMSI Postdoctoral Fellows include demonstrated research ability in statistical and/or applied mathematical sciences, excellent computational skills and the ability to communicate both verbally and in written form. Finally, the preferred applicant will have a strong interest in the SAMSI program areas offered. The deadline for full consideration is December 15, 2016, although later applications will be considered as resources permit.

Please specify which of the two SAMSI research programs you are applying for in your cover letter and why you believe you would be a good fit for SAMSI and the program you choose.

To apply, go to mathjobs.org: SAMSIPD2017 Job #8986

To see these programs visit: www.samsi.info/QMC and www.samsi.info/CLIM

SAMSI is an Affirmative Action/Equal Opportunity employer

FACULTY POSITIONS IN APPLIED MATHEMATICS

Engineering Sciences & Applied Mathematics

The Department of Engineering Sciences and Applied Mathematics at Northwestern University (**http://www.mccormick.northwestern.edu/applied-math/**) invites applications for two full-time, tenure-track faculty positions. Hiring will be focused on the assistant professor level, although exceptional applicants at a more senior level will be considered.

Requirements include a Ph.D. and demonstrated ability to conduct high-impact interdisciplinary research in applied mathematics. Candidates in all areas are encouraged to apply. We are seeking candidates with strong connections to applications who can enhance the breadth of research activities of the department. Duties involve teaching and research.

The application package should include a curriculum vita, a statement of research accomplishments and interests, and a statement of teaching experience and philosophy. Send applications to **https://facultysearch.mccormick.northwestern.edu/apply/index/ODg**.

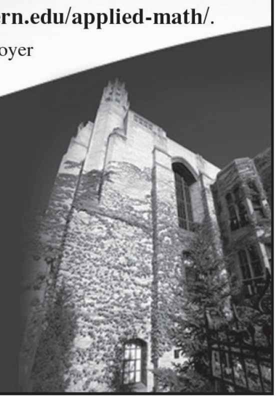
In addition, the applicant should arrange for at least three, but no more than four letters of recommendation. Recommendation letters will be automatically solicited from the letter writers by email after the names are entered in the online application system.

Questions may be sent to **esam-facultysearch@northwestern.edu**
(Subject line: 2017 Faculty Search).

Review of applications will begin November 15, 2016, and will continue until the positions are filled.

For further information, see: **http://www.mccormick.northwestern.edu/applied-math/**.

Northwestern University is an Equal Opportunity, Affirmative Action Employer of all protected classes including veterans and individuals with disabilities. Women and minorities are encouraged to apply. Hiring is contingent upon eligibility to work in the United States.



NORTHWESTERN UNIVERSITY

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES

The *Institute for Computational Engineering and Sciences (ICES)* at *The University of Texas at Austin* is searching for exceptional candidates with expertise in computational science and engineering to fill several Moncrief endowed faculty positions at the Associate Professor level and higher. These endowed positions will provide the resources and environment needed to tackle frontier problems in science and engineering via advanced modeling and simulation. This initiative builds on the world-leading programs at ICES in Computational Science, Engineering, and Mathematics (CSEM), which feature 16 research centers and groups as well as a graduate degree program in CSEM. Candidates are expected to have an exceptional record in interdisciplinary research and evidence of work involving applied mathematics and computational techniques targeting meaningful problems in engineering and science. For more information and application instructions, please visit: www.ices.utexas.edu/moncrief-endowed-positions-app/. This is a security sensitive position. The University of Texas at Austin is an Equal Employment Opportunity/Affirmative Action Employer.

THE UNIVERSITY OF TEXAS AT AUSTIN

Professional Opportunities

Continued from page 10

members of minority groups, and we are strongly committed to increasing faculty diversity. UNCG is an EOE AA/M/F/D/V Employer.

University of North Carolina at Greensboro

Department of Mathematics and Statistics

The University of North Carolina at Greensboro, Department of Mathematics and Statistics seeks applications for one tenure-track position in computational topology at the Assistant Professor rank beginning August 1, 2017. Candidates must hold (or anticipate) a Ph.D. in mathematics or closely-related area by August 1, 2017. Competitive applicants will have research expertise that strengthens our Ph.D. program in computational mathematics. Successful applicants will be expected to excel in teaching, maintain a vigorous research program, seek external research funding, contribute to the interdisciplinary mission at UNCG, and educate a diverse group of undergraduate and graduate students from various backgrounds. Application materials should be submitted electronically to http://https://www.mathjobs.org/jobs/UNCG. Review of applications will begin on December 15, 2016, and will continue until the position is filled.

UNCG is especially proud of the diversity of its student body, which is 43% ethnic minority (http://admissions.uncg.edu/discover-about.php). UNCG has been designated as a Minority-Serving Institution by the U.S. Department of Education. We seek to attract a diverse applicant pool for this position, especially women and members of minority groups, and we are strongly committed to increasing faculty diversity. UNCG is an EOE AA/M/F/D/V Employer.

North Carolina State University

Department of Mathematics

The Mathematics Department at North Carolina State University invites applicants for an endowed, tenured professorship in mathematics. We are seeking exceptionally well-qualified, world-renowned individuals with research interests compatible with those in the department. All areas of pure and applied mathematics will be considered. Candidates must have a Ph.D. in the mathematical sciences, and qualifications for appointment as a tenured associate or full professor. The successful applicant must have an internationally-recognized research program, demonstrated ability to attract external funding, a strong record of mentoring Ph.D. students and post-docs, and a commitment to effective teaching at the undergraduate and graduate levels. The Department of Mathematics has strong research programs in both pure and applied mathematics. Many members of the department participate in interdisciplinary programs and research groups on campus and in the broader Research Triangle community. More information about the department can be found at http://www.math.ncsu.edu.

To submit your application materials, go to http://www.mathjobs.org/jobs/ncsu. For consideration, applicants should submit a vita and a list of four references with complete contact information. You will then be given instructions to go to https://jobs.ncsu.edu/postings/74279 and complete a Faculty Profile for the position.

Write to math-jobs@math.ncsu.edu for questions concerning this position.

NC State University is an Equal Opportunity and Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, national origin, religion, sex, age, veteran status, or disability. In addition, NC State University welcomes all persons without regard to sexual orientation or genetic information. Persons with disabilities requiring accommodations in the application and interview process please call (919) 515-3148. We welcome the opportunity to work with candidates to identify suitable employment opportunities for spouses or partners.

Queen's University

Department of Mathematics and Statistics

The Department of Mathematics and Statistics, faculty of Arts and Science at Queen's University, Kingston, Canada, invites applications for a tenure-track faculty position in applied mathematics and mathematics and engineering at the rank of Assistant Professor with a starting date of July 1, 2017.

In areas related to mathematics and engineering, there are presently prominent research groups in geometric mechanics, control theory, and information and communication theory.

Candidates with expertise in any of the following areas are strongly encouraged to apply: mathematical fluid and continuum mechanics, partial differential equations, numerical analysis, and control of mechanical systems.

For more information about the Mathematics and Engineering program, please see http://www.mast.queensu.ca/meng/. A successful candidate will be expected to work in any of these or complementary research areas, and to contribute to both the graduate and undergraduate programs. A candidate who joins the Mathematics and Engineering group will be expected to obtain a license as a Professional Engineer; an undergraduate degree in engineering is a strong asset towards obtaining the license.

For the full position announcement, please visit: http://www.mast.queensu.ca/positions/.

California Institute of Technology

Department of Computing and Mathematical Sciences

The Department of Computing and Mathematical Sciences (CMS) at the California Institute of Technology (Caltech) invites applications for tenure-track or tenured faculty positions. CMS is a unique environment where innovative, interdisciplinary, and foundational research is conducted in a collegial atmosphere. Candidates in all areas of computing and mathematical sciences are invited to apply, including (but not limited to) learning and computational statistics, security and privacy, networked and distributed systems, optimization and computational mathematics, control and dynamical systems, theory of computation and algorithmic economics, scientific computing, etc. Additionally, we are seeking candidates who have demonstrated strong connections to other fields, including the mathematical, physical, biological, and social sciences.

A commitment to world-class research, high-quality teaching, and mentoring is expected. The initial appointment at the Assistant Professor level is for four years and is contingent upon the completion of a Ph.D. degree in computer science, applied mathematics, or related field.

Applicants are encouraged to have all their application materials on file by October 21, 2016, but applications will be accepted until the end of December. For a list of documents required and full instructions on how to apply online, please visit http://www.cms.caltech.edu/search.

Questions about the application process may be directed to: search@cms.caltech.edu.

Caltech is an Equal Opportunity/Affirmative Action Employer. Women, minorities, veterans, and disabled persons are encouraged to apply.

Donald Knuth Talks Satisfiability and Combinatorics

By James Case

Donald Knuth, professor emeritus in the Department of Computer Science at Stanford University, delivered the John von Neumann Lecture at the 2016 SIAM Annual Meeting, held in Boston this July. Perhaps best known for his invention of the T_EX and *METAFONT* systems for computer typesetting, Knuth has enjoyed a long and fruitful research career while composing a multi-volume treatise on *The Art of Computer Programming*. Much of his lecture was drawn from Volume 4, Fascicle 6 of that still-emerging series.[1]

Knuth spoke about the satisfiability of Boolean formulae $F(x_1,...,x_n)$, expressed as collections of OR clauses separated by ANDs in what is often described as "conjunctive normal form." Is it possible to assign values from the set $\{0,1\}$ to the variables $(x_1,...,x_n)$ in such a way that $F(x_1,...,x_n)=1$? Due to its growing importance, the problem has come to be known as SAT. If $\wedge$ or $\vee$ denote AND and OR respectively, if $\bar{Q}$ denotes the complement of Q in $\{0,1\}$, and if

$$F(x_1,x_2,x_3,x_4)=(x_1\vee x_2\vee\bar{x}_3)$$
$$\wedge(x_2\vee x_3\vee\bar{x}_4)\wedge(x_3\vee x_4\vee x_1)$$
$$\wedge(x_4\vee\bar{x}_1\vee x_2)\wedge(\bar{x}_1\vee\bar{x}_2\vee x_3)$$
$$\wedge(\bar{x}_2\vee\bar{x}_3\vee x_4)\wedge(\bar{x}_3\vee\bar{x}_4\vee\bar{x}_1),$$

one can easily verify that the assignment $(x_1,x_2,x_3,x_4)=(0,1,\#,1)$ satisfies $F(x_1,x_2,x_3,x_4)=1$, where $\#$ may denote either 0 or 1. On the other hand, no such assignment satisfies $G(x_1,x_2,x_3,x_4)=1$ if G is the formula obtained by appending the clause $(\bar{x}_4\vee x_1\vee\bar{x}_2)$ to F. Finally, because G is invariant under the cyclic permutation $x_1\to x_2\to x_3\to x_4\to \bar{x}_1\to\bar{x}_2\to\bar{x}_3\to\bar{x}_4\to x_1$, the omission of *any* clause in G yields a problem equivalent to the satisfiability of F. For this and other reasons, the Boolean formula G is well known to computer scientists.

Applying De Morgan's laws to a formula F in conjunctive normal form yields a formula $\bar{F}$ in "disjunctive normal form," consisting of a collection of AND clauses separated by ORs. So F is unsatisfiable if and only if $F(x_1,...,x_n)\equiv 0$, which happens if and only if $\bar{F}(x_1,...,x_n)\equiv 1$. And since a formula identically equal to 1 is called a *tautology*, one may conclude that a given formula F is satisfiable unless $\bar{F}$ is a tautology.

SAT solvers are algorithms for discovering satisfactory Boolean assignments – or proving that none exist. Such solvers are of growing practical importance, in part because a host of significant combinatorial problems can be reduced to SAT problems, and in part because highly-effective SAT solvers (capable of handling "industrial strength" problems with millions of variables) have suddenly and unexpectedly become available. After all, a polynomial algorithm capable of deciding whether or not a given Boolean formula F is satisfiable would confirm that $P=NP$!

Employing the notation kSAT conveniently indicates that a particular Boolean formula consists of clauses containing no more than k variables. One can show that every SAT problem is logically equivalent to a 3SAT problem whose "size" is polynomial in the size of the original.

Knuth used the early part of his lecture to exhibit a variety of problems, including several about map coloring, reducible to SAT problems. Among the latter were the McGregor map of order ten (see Figure 1) and an ordinary 8×8 chessboard, colored

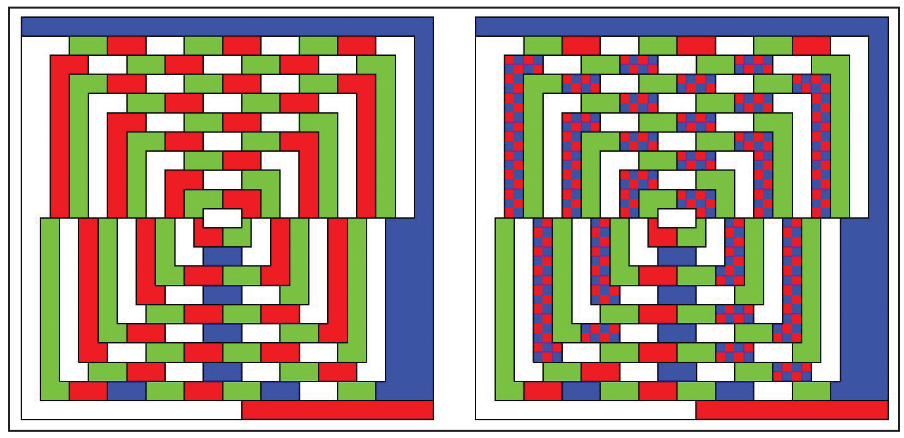


Figure 1. *The McGregor map of order 10 may be 4-colored in at least 2^23= 8,308,608 ways. In one way, one color (blue here) is used only 7 times! Image credit: Donald Knuth and [2].*

in such a way that any two squares separated by a queen's move are the same color.

Questions concerning John Horton Conway's famous *Game of Life* can be similarly resolved. The game is played on a succession of rectangular grids, and each cell is colored either black or white by a solitary player who cannot deviate from the following two-rule strategy:

(i) A black cell shall remain black if and only if either two or three of its eight neighboring cells are currently black

(ii) A white cell shall become black if and only if three of its eight neighbors are currently black.

Knuth displayed initial configurations, all found by SAT solvers, of black and white cells on a 7×15 grid that spell out the word LIFE in block letters five cells high after one, two, or three player moves. He also presented several pairs of configurations in which designated cells alternate between black and white on successive moves (see Figure 2). Finally, he exhibited a configuration known as a "Garden of Eden," which has no predecessor within the game.

To educate himself about SAT solvers, Knuth began building his own several years ago. He has since built at least a dozen of them, each more sophisticated than the last. During his talk, Knuth displayed a progression of these solvers, labeled SAT0 through SAT13. All accept input clauses in the following format,

```
 x2   x3  ~x4
 x1   x2   x4
~x1   x2   x4
~x1  ~x2   x3
~x2  ~x3   x4
~x1  ~x3  ~x4
 x1  ~x2  ~x4
 x1  ~x2   x4
 x1   x2   x3,
```

corresponding to Ronald Rivest's well-known formula $G(x_1,x_2,x_3,x_4)$. Knuth's solvers all return YES or NO answers, along with a satisfactory $0/1$ vector if any such exist. It is important to note, when comparing results, that many authors require their algorithms to return *all possible* satisfactory vectors.

A brute-force SAT solver would construct and search a binary tree with 2^n-1 branch nodes and 2^n leaf nodes, in which all paths from root to leaf are of length n. Each leaf node must be labeled either $F=0$ or $F=1$, depending on whether the formula F is satisfied by the $0/1$ values labeling the edges of the path from root to leaf.

Backtracking methods seek to guess a partial solution to the problem by assigning $0/1$ values to a subset of the variables $x_1,...,x_n$, and testing the hypothesis that those values are correct by extending the partial solution to a complete one. Success confirms the hypothesis while failure rejects it, indicating only that the initial guess was wrong and another must be made. Therein lies the rub, for there are C^k_n k-subsets of $N=\{1,...,n\}$, and 2^k ways of assigning $0/1$ values to a k-subset of the n problem variables.

It can be obvious which subsets to choose and which values to assign. If a single variable x_i appears in every clause, the set $\{x_i\}$ and value $x_i=1$ are clear choices. Or if x_i appears in "almost all" of the clauses, then the previous guess still seems appropriate. The choice $x_i=1$ then eliminates (i.e. satisfies) all but a few of the clauses, and $\bar{x}_i$ may be stricken from those that remain, leaving behind a greatly-reduced SAT problem. There is, of course, no guarantee that the latter will prove solvable, but its solvability remains a natural hypothesis to explore. Advanced SAT solvers employ a variety of strategies—some of them are quite elaborate—when deciding which subsets to choose and what values to assign.

Knuth referenced a 1982 paper [1] that presented empirical evidence indicating that 0-level algorithms outperform 1-level algorithms on fairly small problems, while 1-level algorithms excel both 0-level and 2-level algorithms on slightly larger problems. The same two authors had previously obtained asymptotic estimates indicating that truly large problems will require higher-level algorithms. As matters stand, art rather than science indicates which algorithms are most appropriately applied to which problems.

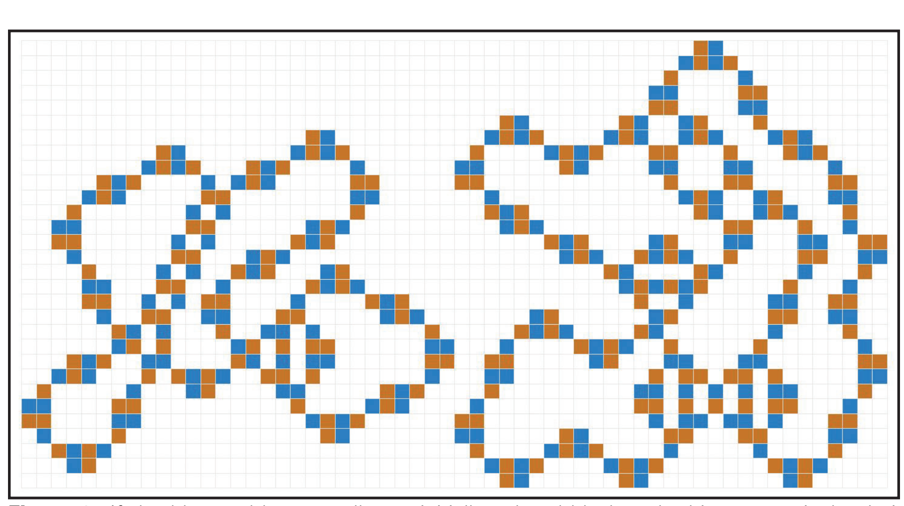


Figure 2. *If the blue and brown cells are initially colored black and white respectively, their colors will alternate on subsequent moves. Image credit: Donald Knuth and [2].*

Having exercised several of his SAT solvers on the following puzzle: *Find a binary sequence $x_1,...,x_8$ containing no three equally-spaced 0s and no three equally-spaced 1s,* Knuth illustrated a number of the results with rows of red and blue balls rather than 0s and 1s. There are six admissible sequences, none of which is extendable to an admissible sequence of length 9. This is a special case of the general fact that, given any two positive integers i and j, along with sufficiently-large n, every binary sequence $x_1,...,x_n$ must contain exactly i equally-spaced 0s and exactly j equally-spaced 1s. The smallest such n is denoted $W(i,j)$ in honor of B. L. van der Waerden, who proved an even more general result. Despite the best efforts of current SAT solvers, relatively little is known about the numbers $W(i,j)$.

The most promising SAT solving applications appear to lie in the fields of hardware and software verification. One can usually represent a typical design as a relation R on $S\times S$, where S denotes a set of admissible "state vectors" $X=(x_1,...,x_n)$ and X' is a possible successor of X if and only if $(X,X')\in R$. Thus, a sequence $X_0\to X_1\to\cdots\to X_\omega$ is possible if and only if $(X_0,X_1)\in R\,\&\,(X_1,X_2)\in R\,\&\,...\,\&(X_{\omega-1},X_\omega)\in R$. One would like to know, for instance, if there are any sequences in which X_0 is an initial state and X_ω generates an error message.

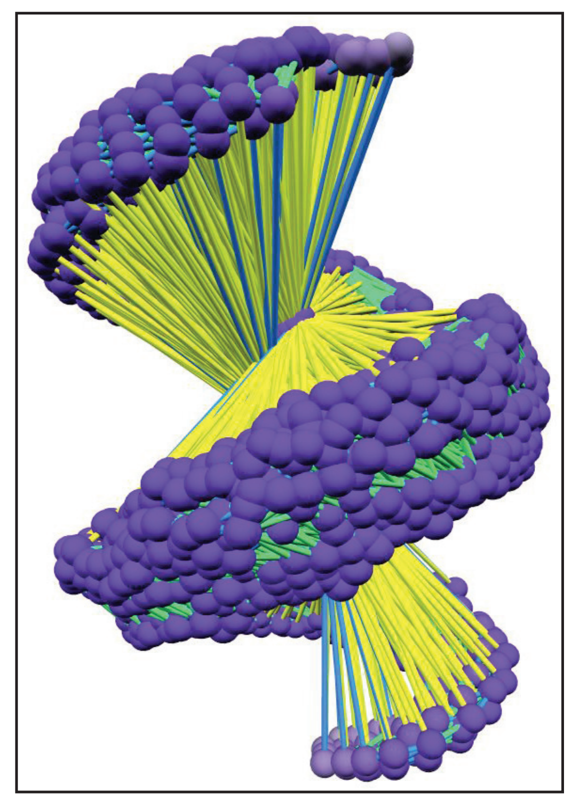


Figure 3. *The purple balls represent Boolean variables, while the multicolored connecting rods identify pairs of variables that appear together in one, few, or many OR clauses. Image credit: Carsten Sinz and [2].*

In the last part of his lecture, Knuth displayed a series of eye-catching illustrations (see Figure 3) dramatizing the complexity of the problems that current SAT solvers are able to resolve. In some of his images, differently-colored connectors indicate the number of clauses containing both connected variables. Such information is invaluable to backtrackers deciding which sets to try and what values to assign.

The lecture was well attended, and not surprisingly, the audience seemed quite spellbound by the progress Knuth reported.

References

[1] Brown, C.A., & Purdom, P.W. (1982, May). An Empirical Comparison of Backtracking Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 4*(3), 309-316.

[2] Knuth, D.E. (2015). *The Art of Computer Programming. (Vol. 4, Fascicle 6: Satisfiability).* Boston, MA: Addison-Wesley.

James Case writes from Baltimore, Maryland.

[1] http://www-cs-faculty.stanford.edu/~uno/taocp.html