Tackling Research Debt

In the book *Indiscrete Thoughts*, Gian-Carlo Rota wrote an essay titled “Problem Solvers and Theorizers,” in which he divided up mathematicians into two types (Rota, 2010). Problem-solvers identify unsolved problems, then, with lazer-focus, attempt to be the first person to solve that problem. Theorizers, on the other hand, take already-solved problems, and try to trivialize them, to make them obvious. In other words, the theorizer’s goal is to shed conceptual or philosophical light on a problem such that the solution becomes obvious once the problem is observed in the right way.

Over the last century, there has been a heavy shift, both in science and mathematics, away from theorizing towards problem-solving. For instance, most mathematical journals will not accept new proofs for already-proved theorems. The problem with this approach is that, while it solves more problems in the short term, it makes it more difficult to solve problems in the long term, as researchers must learn more complex systems than necessary (since there are fewer people around to trivialize them).

The *Distill* group (`https://distill.pub/`) has recognized this fact, and has given it a term—Research Debt. Essentially, you can think of research as having two components—making discoveries and making discoveries understandable. Making discoveries *without* making them understandable incurs “debt” within the whole system. Just like normal debt, if the debt is unpaid, the interest payments show up in the fact that each researcher has more work to do to understand their field (Olah and Carter, 2017).

The Distill group is focused specifically on machine learning, and has organized both a journal and a prize. The journal is an *interactive* academic journal. That is, they go beyond typical PDFs that most online journals provide, and actually include simulations and interactive visualizations in their publications.

The prize is $10,000 USD, and is given for researches in machine learning who, essentially, clarify things that we formally know but don’t intuitively understand, or for reframing problems in a much easier-to-understand way. Nominations for the prize should be send to `prize@distill.pub`.

New Unified Model of Specified Complexity Formulated

One of the problems within the Intelligent Design research community is a plethora of definitions and terms for very similar things. Functional Information (Hazen et al., 2007), Functional Sequence Complexity (Durston et al., 2007), Algorithmic Specified Complexity (Ewert, Dembski, and Marks II, 2014), Algorithmic Mutual Information (Milosavljevi, 1995), and other models all use fundamentally similar mathematics, but the profusion of models has made systematization difficult.

In the latest *Bio-Complexity*, George Montañez creates a unified mathematical model that can be used for all of these concepts (Montañez, 2018). The model is

$$SC(x) = -\log_2\left(r\frac{p(x)}{\nu(x)}\right).$$

Here, $p(\nu)$ is the probability distribution of x, $\nu(x)$ is a specification function (either discrete or continuous but uniformly positive), and r is a scaling factor where $r \geq \nu(\Omega)$ to normalize the distribution. The $-\log_2()$ is applied to convert the probability into bits. Additionally, a signficance level α can be applied by adding $\log_2(\alpha)$ to the result.

Montañez notes that specified complexity can lose some efficiency compared to precise probabilities, but that the benefit is that the calculation can be more easily utilized when precise probabilities are not known. The primary requirement for the application of Specified Complexity is the detachment principle, which states that the specification function must be determined prior to observing the data.

For those with an interest in Specified Complexity, in 2016 The Blyth Institute put out a video tutorial on the mathematics of Specified Complexity (focusing on Algorithmic Specified Complexity), available at `https://youtu.be/5CWu_8CTdDY`. This obviously does not include the generalization found in the presently-discussed paper, but might provide a decent background for understanding the concepts within it.

Blyth Institute Develops Improved Notation for Higher-Order Derivatives

It is commonly known that, in standard calculus, while first-order differentials (such as $\mathrm{d}y$ and $\mathrm{d}x$) can generally be manipulated algebraically, second-order and higher differentials (such as d^2y and $\mathrm{d}x^2$) cannot be so manipulated. In order to manipulate them, special formulas such as Faà di Bruno's formula must be used (the better-known special case of this formula is the chain rule for the second derivative). That is, one cannot just multiply and divide by higher-order differentials and expect to get back consistent and correct results.

This surprising fact is generally left unexplained in standard textbooks. Blyth Institute director Jonathan Bartlett discovered that the reason for this stretches back well over a hundred years, when the formula for the second derivative ($\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$) was originally established. It turns out that a change in this notation will render the differentials for the second derivative to be fully algebraically manipulable, with no apparent drawbacks.

The new notation, interestingly, was developed by using standard calculus rules for taking the derivative, and simply applying those same rules to the first derivative (which is already algebraically manipulable). If you take the first derivative seriously as a quotient (since $\frac{\mathrm{d}y}{\mathrm{d}x}$ is already in the form of a quotient), then simply applying the quotient rule and simplifying yields the improved formula for the second derivative.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}^2x}{\mathrm{d}x^2}$$

Faà di Bruno's formula (and the chain rule for the second derivative) is still valuable because it tells you *how* to manipulate the second derivative in certain cases (which is not altogether obvious), but it is no longer needed to justify the manipulations themselves, which are now just simple algebraic transformations.

This was recently published in the paper "Extending the Algebraic Manipulability of Differentials," published in the journal *Dynamics of Continuous, Discrete and Impulsive Systems, Series A: Mathematical Analysis* (Bartlett and Khurshudyan, 2019a). The paper lays the groundwork for future investigations into the nature of the differential.

Ara Avetisyan and Asatur Zh. Khurshudyan Win Prestigious Science Award

This year, Ara Avetisyan and Asatur Zh. Khurshudyan received the *President Prize of the Republic of Armenia* in the area of "Technical Sciences and Information Technologies." This award is issued by the republic of Armenia for significant contributions to diverse areas of science and mathematics. Avetisyan and Khurshudyan won the award for their work on using Green's functions in control systems analysis, specifically for the monograph *Controllability of Dynamic Systems: The Green's Function Approach* (2018). The official presentation of the award will be on June 10.

The Blyth Institute has leaned on Dr. Khurshudyan for mathematical help for several projects, including the recently published papers "Extending the Algebraic Manipulability of Differentials" (Bartlett and Khurshudyan, 2019a) and "Numberphile's Proof for the Sum $1 + 2 + 3 + \ldots$" (Bartlett and Khurshudyan, 2019b), as well as others which are still pending publication. The Blyth Institute also participated in developing the monograph for which the award was made, with director Jonathan Bartlett's acting as the monograph's editor.

Dr. Khurshudyan said that he plans to use his part of the prize money to help fund scholarships for students in Armenia.

Engineering Working Group Launches

For the past year, an informal group of engineers across many disciplines, together with biologists, scientists, medical researchers, and technologists, have been meeting together to explore the intersection between engineering and biology. The group hopes to enable a greater understanding of biological systems, generate better predictions of future discoveries, empower more productive research, and better understand the causal requirements underlying biological design. Toward this end, a number of projects are either underway or in development, including a survey of engineering in the biology literature, detailed engineering models of the bacterial flagellum (and developing requisite reusable modeling tools), a survey of repeatable and reversable adaptive mechanisms in biology, developing coherence metrics for engineered artifacts, and classification of control mechanisms in living systems. The group is planning to organize more formally in the near future, and will make announcements to this effect in the coming months.

While this is not a Blyth Institute effort, anyone who would like to participate in this project can contact The Blyth Institute for more information.

Avetisyan, A S and A Zh Khurshudyan (2018). *Controllability of Dynamic Systems: The Green's Function Approach*. Newcastle: Cambridge Scholars Publishing.

Bartlett, J L and A Zh Khurshudyan (2019a). "Extending the Algebraic Manipulability of Differentials". In: *Dynamics of Continuous, Discrete and Impulsive Systems, Series A: Mathematical Analysis* 26.3, pp. 217–230.

Bartlett, J L and A Zh Khurshudyan (2019b). "Numberphile's Proof for the Sum $1 + 2 + 3 + \ldots$" In: *Communications of the Blyth Institute* 1.1, pp. 54–55.

Durston, K K et al. (2007). "Measuring the functional sequence complexity of proteins". In: *Theoretical Biology and Medical Modelling* 4.47. DOI: 10.1186/1742-4682-4-47.

Ewert, W, W A Dembski, and R J Marks II (2014). "Algorithmic Specified Complexity". In: *Engineering and the Ultimate: An Interdisciplinary Investigation of Order and Design in Nature and Craft*. Ed. by J Bartlett, D Halsmer, and M R Hall. Blyth Institute Press, pp. 131–151. DOI: 10.33014/isbn.0975283863.7.

Hazen, R M et al. (2007). "Functional information and the emergence of biocomplexity". In: *PNAS* 104.suppl 1, pp. 8574–8581. DOI: 10.1073/pnas.0701744104.

Milosavljevi, A (1995). "Discovering dependencies via algorithmic mutual information". In: *Machine Learning*, pp. 35–50.

Montañez, G D (2018). "A Unified Model of Complex Specified Information". In: *Bio-Complexity* 2018.4.

Olah, C and S Carter (2017). "Research Debt". In: *Distill*. URL: https://distill.pub/2017/research-debt.

Rota, G C (2010). "Problem Solvers and Theorizers". In: *Indiscrete Thoughts*, pp. 45–46.