# A Fresh Look at Boolean Functions: Progress Report

Thomas Courtade, Pulkit Grover, Madars Virza, and Gowtham Kumar

The team is working on three main problems, each of which are listed below.

## I. Which Boolean functions are most informative?

This topic is inspired by the following conjecture of Kumar and Courtade:

**Conjecture 1:** *Let $X^n$ be i.i.d. Bernoulli(1/2), and let $Y^n$ be the result of passing $X^n$ through a memoryless binary symmetric channel with crossover probability $\alpha$. For any Boolean function $b : \{0,1\}^n \to \{0,1\}$, we have*

$$I(b(X^n);Y^n) \leq 1 - H(\alpha). \tag{1}$$

At first sight, Conjecture 1 appears to be a simple consequence of known data processing inequalities. However, over the course of our investigation we learned that, despite its apparent simplicity, standard information-theoretical manipulations appear incapable of establishing (1). To us, Conjecture 1 represents the simplest, nontrivial embodiment of Boolean functions in an information-theoretic context. In words, Conjecture 1 asks: "*What is the most significant bit that $X^n$ can provide about $Y^n$?*"

Despite their fundamental roles in computer science and digital computation, Boolean functions have received relatively little attention from the information theory community. The recent work [KKBS] is perhaps most relevant to our Conjecture 1 and provides compelling motivation for its study. In [KKBS], the authors prove that for $n$ and $\Pr\{b(X^n)=0\} \geq 1/2$ fixed, $I(b(X^n);X_1)$ is maximized by functions $b$ which satisfy $b(X^n)=0$ whenever $X_1 = 0$ (i.e., when $b$ is *canalizing* in $X_1$). The motivation for considering this problem comes from computational biology, where Boolean networks are used to model dependencies in various regulatory networks. More details can be found in [KKBS; SJ] and the references therein.

Conjecture 1 is also related to the *Information Bottleneck Method* [TPB], which attempts to solve the optimization problem

$$\min_{p(u|x^n)} I(X^n;U) - \lambda I(Y^n;U). \tag{2}$$

For a given $\lambda > 0$, the optimizing $U$ is purportedly the best tradeoff between the accuracy of describing $Y^n$ and the descriptive complexity of $U$. In our setting, $b(X^n)$ plays the role of $U$, and we constrain the descriptive complexity to be at most one bit.

A more concrete example comes in the context of gambling. To this end, suppose $Y^n$ is a simple model for a market of $n$ stocks, where each stock doubles in value or goes bankrupt with probability 1/2, independent of all other stocks. If an oracle has access to side information $X^n$, and we are allowed to ask *one* yes/no question of the oracle, which question should we ask to maximize the rate at which our wealth grows? The validity of Conjecture 1 would imply that we should only concern ourselves with the performance of a single stock; say $Y_1$. This is readily seen as a consequence of known results on gambling with side information [CT, Theorem 6.2.1], since putting $b(X^n) = X_1$ yields

$$I(b(X^n);Y^n) = I(X_1;Y^n) = I(X_1;Y_1) = 1 - H(\alpha), \tag{3}$$

hence the conjectured upper bound (1) is attainable and represents the maximum possible increase in doubling rate.

Finally, we point out that (1) is related in spirit to the notion of *average sensitivity* of Boolean functions. This topic has received a great deal of attention in the computer science literature (cf. [O'D]). To see the connection to sensitivity, note that (1) can be rewritten as

$$H(b(X^n)|Y^n) \geq H(b(X^n)) - 1 + H(\alpha). \tag{4}$$

For fixed $\Pr\{b(X^n)=0\}$, the right hand side of (4) is constant. Hence, the conjecture essentially lower bounds the output uncertainty of Boolean functions with respect to noisy inputs.

*Progress thus far*

We have written a paper which was accepted to and presented at ISIT. This paper is attached to the report. Detailed statements of the results can be found therein. However, in a nutshell:

**Main Results:** We have proven that it suffices to consider a very small set of functions to prove Conjecture 1. Indeed, this allows us to prove the conjecture for $n = 7$ (which would require enumerating $3.4 \times 10^{38}$ functions if done exhaustively). Second, we have an algorithmic proof of a secondary Conjecture (also of significant interest).

In addition to the results included in the ISIT paper, we have successfully proven the weaker inequality

$$\sum_{i=1}^{n} I(b(X^n); Y_i) \leq 1 - H(\alpha) \tag{5}$$

using Fourier-analytic techniques similar to those employed in [KKBS]. However, this Fourier-analytic approach appears incapable of establishing the stronger statement of Conjecture 1.

**Follow on work by ourselves and others:** This conjecture was well-received by the information theory community. Already, several researchers have devoted some effort to it. Most notably, Sudeep Kamath at UC Berkeley and Chandra Nair at CUHK. We are currently in the process of writing a journal paper on our results.

## II. Information-friction and the required energy to compute a function

What are the fundamental limits on the required energy of computing? The question has been investigated at least since 1960s (if not earlier) in the attempt of understanding a theoretical physics paradox: the Maxwellian demon [LR]. However, the analysis in physics neglects two important phenomena that real-world computation encounters: frictional losses, and noise.

The goal here is to obtain some understanding of energy required for computing where friction and/or noise are explicitly accounted for. The simplest computation problem, using binary inputs, is perhaps computation of a boolean function. After all, every binary function can be represented as a vectorization of boolean functions!

In our attempt to address the issue, we began with a simple and yet relevant computational problem: the problem of computing *majority* of input bits, which applicable to estimating a channel input that is sent across a channel using a repetition strategy (i.e. 0, is repeated many times, e.g. 00000). The essential question here is:

What is the tradeoff between energy of decoding (computing a boolean function) at the decoder and the error probability? What can be said about a more general (non-repetition) code?

*Results and work-in-progress*

**Implementation and energy model**: In standard models of computation (e.g. the Turing machine model, and even Yao's communication complexity model) and circuits, the energy for moving information is often ignored. In our earlier experimental circuit observations for decoding [GGR], the energy consumed by circuit wires is significant. This motivated us to formulate a model of "Information-Friction," illustrated in Fig. 1.

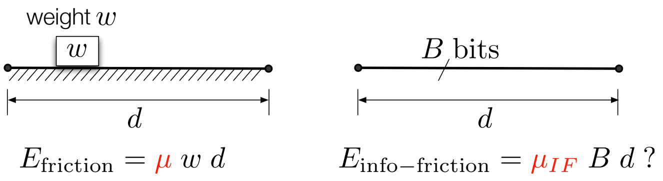
weight $w$
$w$
$d$
$E_{\text{friction}} = \mu\, w\, d$
$B$ bits
$d$
$E_{\text{info-friction}} = \mu_{IF}\, B\, d$ ?

Fig. 1. The model for energy in information-friction is analogous to classical friction. Instead of moving a mass, a "mass of information" is moved from one point to the other. The information-friction model applies to communication across circuit wires, optical fibers, and even moving a stored memory physically from one point to the other. It also applies to biological communication, e.g. through the axon on a neuron.

**Definition 1 (The** bit-meters **cost of a link and of a circuit):** The bit-meters cost of a *link* in a computation Comp is the product of the number of bits carried on that link and the Euclidean distance between the nodes at the ends of the link. The bit-meters for the entire circuit Ckt is the sum of bit-meters for all the links in Comp.

The energy model that we propose is a model based on "Informational Friction":

**Definition 2 (Energy consumed by informational friction in a computation):** The coefficient of informational friction, denoted by $\mu$, characterizes the energy required for computation in our model. This energy is given by $E = \mu \times \mathrm{bm}$, where $\mathrm{bm}$ is the $\mathrm{bit\text{-}meters}$ consumed in the given circuit and computation.

Our theoretical results provide the first such tradeoffs on energy required for decoding repetition codes, and demonstrate that asymptotically (in the limit of $P_e \to 0$, where $P_e$ is the target error probability), *the energy consumed in wires (i.e., moving information) dominates that consumed in the computational nodes*.

The following result is a simple observation made in [5].

**Theorem 1:** The number of $\mathrm{bit\text{-}meters}$ required for decoding one bit in repetition coding for a repetition code transmitted across a BSC with crossover probability $p_{ch}$ is at least:

$$\mathrm{bit\text{-}meters}\ \textsf{per-bit}(\mathrm{repetition}) \geq \Omega\left(\mu \frac{\log \frac{1}{P_e}}{\log \frac{1}{p_{ch}}}\right). \qquad (6)$$

We have been able to extend this understanding to encoding/decoding of an error correcting code. Our bounds [5] hold for all codes and all decoding algorithms that achieve a target rate and a target error probability.

More recently, we have extended this understanding to incorporate time-constrained computing (of, e.g., a Fourier transform) [6]. Implications of these results in neuronal networks is under progress. Our hope is to be able to provide a fundamental understanding (including relevant tradeoffs) of computation in the brain that incorporates frictional losses as well as noise in links.

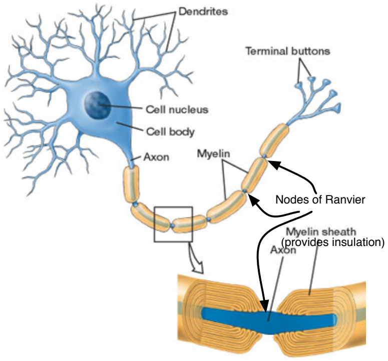

Fig. 2. Energy requirements for communication along the axon of a neuron can also be modeled using information friction: the "nodes of Ranvier" act as relays, regenerating spikes and transferring information.

## III. The problem of verifying computation

A difficulty that arises when studying the efficiency of proofs for *arbitrary* $\mathsf{NP}$ statements is the problem of *representation*. Proof systems, such as [BCI$^+$], are typically designed for inconvenient $\mathsf{NP}$-complete languages such as circuit satisfiability or algebraic constraint satisfaction problems, while in practice, many of the problem statements we are interested in proving are easiest to express via algorithms written in some high-level programming language. Until all existing reductions from programs written in high-level languages to Boolean circuits, such as FairPlay [BDNP], have had quadratic overhead, making them completely impractical for all, but smallest workloads.

Modern compilers can efficiently transform algorithms into a program to be executed on a *random-access machine* (RAM). Therefore, we seek proof systems that efficiently support $\mathsf{NP}$ statements expressed as the correct execution of a RAM program. Our computationally sound proof system directly supports the $\mathsf{NP}$-complete language consisting of correct executions of programs on TinyRAM, a small, nondeterministic random-access machine that we designed.

To achieve our goal of efficient reduction from TinyRAM to arithmetic circuits, we use the framework of [BSCGT] and are very interested in constructing minimal circuits for verifying correct execution of a single TinyRAM time-step and verifying correctness of memory accesses.

One of our main contributions are transition function and memory checker of novel design that both make heavy use of non-determinism and the algebraic structure of the field of our arithmetic circuit. This allows us to achieve our result: the first system that transforms arbitrary C programs in asymptotically optimal $\tilde{O}(T)$-sized arithmetic circuits and prove their correct execution in zero-knowledge [BSCG$^+$].

The work was presented at CRYPTO 2013 [7].

## IV. Budget and Expenditures

Primary expenditures were related to conference travel and registration. A detailed breakdown is as follows: \$745.64 was spent to attend ITA at UCSD and \$2749.46 was spent to attend ISIT. The remaining funds (\$1504.90) will be used by Pulkit to cover his ISIT expenses.

## V. Team Meetings

- Meetings take place nearly daily between Gowtham and Tom. This culminated in their ISIT paper.
- 12/2-12/5: NSF Site visit at Purdue. Both Madars and Tom attended. They discussed problems B and C. Posters were presented on each problem.
- 2/10-2/15: ITA 2013 at UCSD. Pulkit, Tom and Gowtham had a discussion. Gowtham and Tom presented a poster.
- There were multiple rounds of emails exchanged between the group in late September-October. The topic of superlinear lower bounds on circuit complexity was discussed, and we realized it is a really really hard problem (mostly as a result of Madars' patient explanations).
- Team meeting (Gowtham, Tom, Pulkit) at ISIT'13 conference
- Team meeting (Madars & Tom) at CSoI NSF site visit / poster presentations.
- Several skype meetings with Madars, Pulkit, Tom, roughly once every two months.

## VI. Problems/hurdles encountered

Despite the good productivity, it has been extremely challenging to encourage collaboration amongst people in different locations. This is mostly due to the fact that everyone is busy with various things and it is easy to put off team meetings. Nevertheless, we enjoyed our interactions, and having this seed grant did encourage these interactions.

Because we could not arrive at a topic of immense interest to all involved, we have decided to not request a renewal of this grant. We do thank CSoI for providing us this wonderful platform to interact on.

## VII. Outcomes: Conference and Poster Presentations

[1] G. Kumar and T. Courtade, "Which Boolean Functions are Most Informative?," 2013 International Symposium on Information Theory, Istanbul, Turkey, July 8-12, 2013 (also available online: *arXiv:1302.2512* [cs.IT]).

[2] G. Kumar and T. Courtade, "Which Boolean Functions are Most Informative?," Poster presentation at the 2013 Information Theory and Applications Workshop (ITA) at UCSD.

[3] G. Kumar and T. Courtade, "Maximally Informative Boolean Functions," Poster presentation at the December, 2012 NSF site visit at Purdue University.

[4] Karthik Ganesan, Pulkit Grover, and Andrea Goldsmith, "How far are LDPC codes from fundamental limits on total power consumption?" 49th Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, Oct. 2012.

[5] P. Grover, " 'Information-Friction' and its impact on minimum energy per communicated bit," 2013 International Symposium on Information Theory, Istanbul, Turkey, July 8-12, 2013.

[6] P. Grover, "Faster Computing Requires Moving Information to Larger Distances" in preparation.

[7] E. Ben-Sasson, A. Chiesa, D. Genkin, E. Tromer and M. Virza. "SNARKs for C: Verifying Program Executions Succinctly and in Zero Knowledge." In proceedings of the 33rd International Cryptology Conference (CRYPTO 2013)

## References

[BCI$^+$] Nir Bitansky, Alessandro Chiesa, Yuval Ishai, Rafail Ostrovsky, and Omer Paneth. Succinct non-interactive arguments via linear interactive proofs. In *TCC*, pages 315–333, 2013.

[BDNP] Assaf Ben-David, Noam Nisan, and Benny Pinkas. Fairplaymp: a system for secure multi-party computation. In *ACM Conference on Computer and Communications Security*, pages 257–266, 2008.

[BSCG$^+$] Eli Ben-Sasson, Alessandro Chiesa, Daniel Genkin, Eran Tromer, and Madars Virza. SNARKs for C: Verifying Program Executions Succinctly and in Zero Knowledge. In *under submission to CRYPTO 2013*, 2013.

[BSCGT] Eli Ben-Sasson, Alessandro Chiesa, Daniel Genkin, and Eran Tromer. Fast reductions from rams to delegatable succinct constraint satisfaction problems: extended abstract. In *ITCS*, pages 401–414, 2013.

[CT] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2nd edition, 2006.

[GGR] Karthik Ganesan, Pulkit Grover, and Jan M Rabaey. The power cost of overdesigning codes. In *SiPS*, October 2011.

[KKBS] Johannes Georg Klotz, David Kracht, Martin Bossert, and Steffen Schober. Canalizing boolean functions maximize the mutual information. *arxiv:1207.7193*, 2012.

[LR] H S Leff and A F Rex. *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*. Insitutite of Physics, 2003.

[O'D] Ryan O'Donnell. Some topics in analysis of boolean functions. In *Proc. STOC '08*, pages 569–578, New York, NY, USA, 2008. ACM.

[SJ] Areejit Samal and Sanjay Jain. The regulatory network of e. coli metabolism as a boolean dynamical system exhibits both homeostasis and flexibility of response. *BMC Systems Biology*, 2(1):21, 2008.

[TPB] N. Tishby, F. C. Pereira, and W. Bialek. The information bottleneck method. In *The 37th Annual Allerton Conference on Communication, Control, and Computing*, pages 368 – 377, September 1999.