Combinational Logic Design in Synthetic Biology

Douglas Densmore
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA 94720-1770
densmore@eecs.berkeley.edu

J. Christopher Anderson
Department of Bioengineering
University of California, Berkeley
Berkeley, CA 94720
jcanderson@berkeley.edu

***Abstract*— Combinational logic circuits are a fundamental building block in today's digital electronics. Combinational logic representations are highly amenable to various levels of abstraction, and manipulation is naturally performed via Boolean algebraic expressions. These properties have allowed automatic VLSI synthesis and analysis capabilities to reach a high level of maturity. In the growing field of Synthetic Biology, devices are emerging which are being classified as "combinational logic". These devices can be constructed using standardized design practices which pave the way to more advanced analysis and construction techniques. However there exist a number of barriers to the application of VLSI design techniques to biological systems. We describe two key barriers and begin the process of examining how difficult they will be to overcome.**

I. Introduction

Digital electronics fundamentally operate using the underlying concept of a binary switch. These switches allow for the expression of Boolean algebraic expressions which can perform mathematical operations and can compose extremely complex decision making systems. The grouping of these switches creates two systems of digital logic: combinational and sequential. Both systems have the notion of binary input signals and binary output signals. They differ in that combinational logic output is strictly a function of the current inputs, whereas sequential logic output depends on the current inputs and the history of those inputs (this is called the "state"). Sequential logic primarily functions as memory elements. Combinational logic, broken down into building blocks called logic gates, is the focus of this paper.

There are a number of basic combinational logic gates including AND, OR, and XOR. The addition of inversion with a NOT gate, leads to NAND, NOR, and XNOR gates as well. Each gate's behavior is described by its truth table (a canonical, tabular representation of the output given a set of inputs). Every Boolean function's truth table has a canonical form called its Sum-of-Products (SOP) expression. This is called a "two level logic expression" where the products (conjunction) of binary variables (called "cubes") are summed (disjunction). Also of importance is that NAND and NOR gates are functionally complete. This means that using only NAND or NOR any other logic expression can be created. The notion of functional completeness and canonical Boolean forms create the basis for much of digital logic synthesis. A key property of combinational logic in digital designs is its ability to support abstraction. Figure 1 illustrates three key abstractions. The first level is the physical layout of a CMOS device. The second level is a schematic of PMOS and NMOS transistors. The final level is the iconography which illustrates a logic gate as a function of its primary inputs. We will illustrate that these abstractions may be difficult to maintain as cleanly in synthetic biology.

Synthetic biology itself is a ground-up approach to genetic engineering wherein DNA molecules are added to cells to introduce new biological functions. The introduction of large DNAs is in many ways similar to the blank slate of a silicon wafer and an analogy to VLSI is attractive. The technology to construct large DNAs and install them into cells exists. This technology is being used to construct a wide variety of useful organisms primarily for chemical production, healthcare, and bioenergy. Additionally, initial efforts are underway to encode combinational logic within a cell. The primary difference between a living system and a silicon wafer is the chemical complexity of the system. Whereas the initial substrate for a semiconductor device is made purely of one element, a cell contains dozens of elements combined into thousands of ornate organic and inorganic compounds which change in composition and concentration over time. This complexity complicates our ability to readily adapt scalable logic devices within a cell.

Much attention has been given recently to encoding combinational logic in cells. Several issues have appeared in the early implementations as illustrated in Figure 1. The same iconography present in electrical engineering is being used for biological systems. Biologists have examined building combinational logic using transcriptional activators [1] [2] [3], synthetic peptides [4], engineered ribsomes [5], and small RNAs [6]. This begs two questions. One is: "What exactly is this logic to be used for?" Some envision the use of cells as massively parallel computation systems. Others envision a more limited role where cells make decisions such as whether to produce a biofuel or deliver a therapeutic agent. The

978-1-4244-3828-0/09/$25.00 ©2009 IEEE

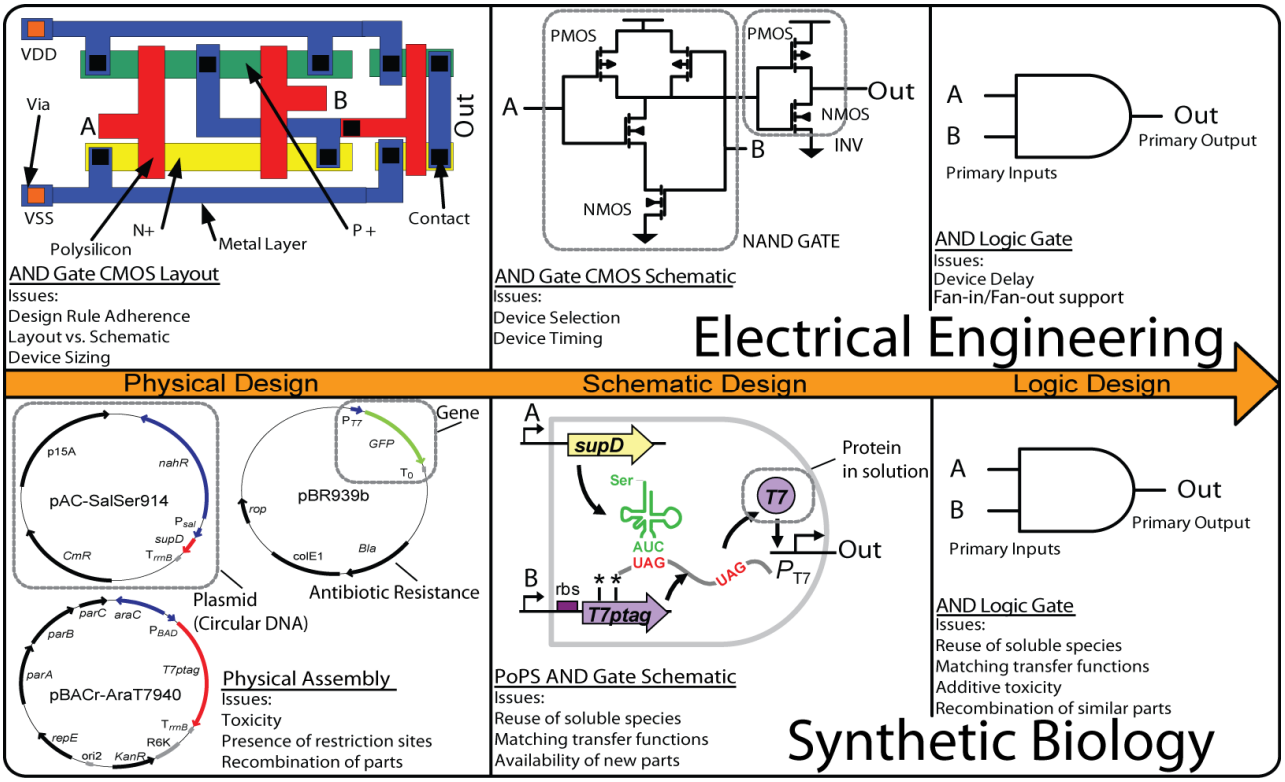

Figure 1: Representations and Abstraction in Combinational Logic

specific application of logic in the context of biological systems is not fully clear, but a scalable logic paradigm is broadly viewed as a highly valuable target in biological engineering. This leads to the second question, "What abstractions hold from the digital world"? This answer illustrates some key fundamental differences between cells and VLSI devices.

This paper focuses on two barriers to the broad application of combinational logic manipulation techniques employed by VLSI tools to biological circuits. These are structural duplication and carrier signal selection.

II. Structural Duplication In Biological Circuits

One key distinction between cells and VLSI devices is the absence of the physical substratum and spatial addressing in a cell. It is true that the DNA molecule itself has distinct physical 1D footprint onto which logical devices might be installed—a gene in a cell does have a distinct physical location on the DNA. However, all biological logic devices thus far described involve the actions of soluble components encoded by the DNA that act through biochemical interactions. The mechanisms for producing these components (transcription and usually also translation) themselves involve biochemical components that are not associated with the DNA; the physical linkage between these soluble components of the device and the DNA components of the device is broken. Therefore biological devices (as they currently are designed) cannot be duplicated within a cell without in effect "shorting each other out". In Figure 2, two very basic digital Boolean manipulations are shown. The biological versions are shown below. Throughout this paper, elements "on the line" represent DNA coding elements whereas those "off the line" are soluble biochemical species encoded by the DNA. In Figure 2's AND gate, two transcription inputs drive the production of an amber suppressor tRNA and the mRNA encode a transcriptional activator with two amber stop codons. When both inputs are activated, the amber stop codons are suppressed during translation resulting in a functional activator. This activator drives transcription from the output promoter. The biological reality illustrates the production of proteins from one coding sequence repressing the promoters for other coding sequences. This will create cross reactivity. This is quite different from a VLSI circuit where one builds a single unique device and then repeats it millions of times on the same die. Implementing such a scalable logic within a cell would require either a different way of physically compartmentalizing the soluble components, or the development of a method for biochemically distinguishing each device from the others. In this latter approach, physical location or wiring connections are replaced by the specificity of biochemical interactions.

III. Carrier Signal Selection In Biological Circuits

The second challenge in constructing biological devices is the diversity of signal carriers within a cell. In VLSI, it is self evident that the information will be channeled throughout the device via the flow of electrons or holes, and the theory of how to describe this flow in the form of voltages and currents is well understood. In biology, there are far more choices as to what the signal carrier might be: there can be electron

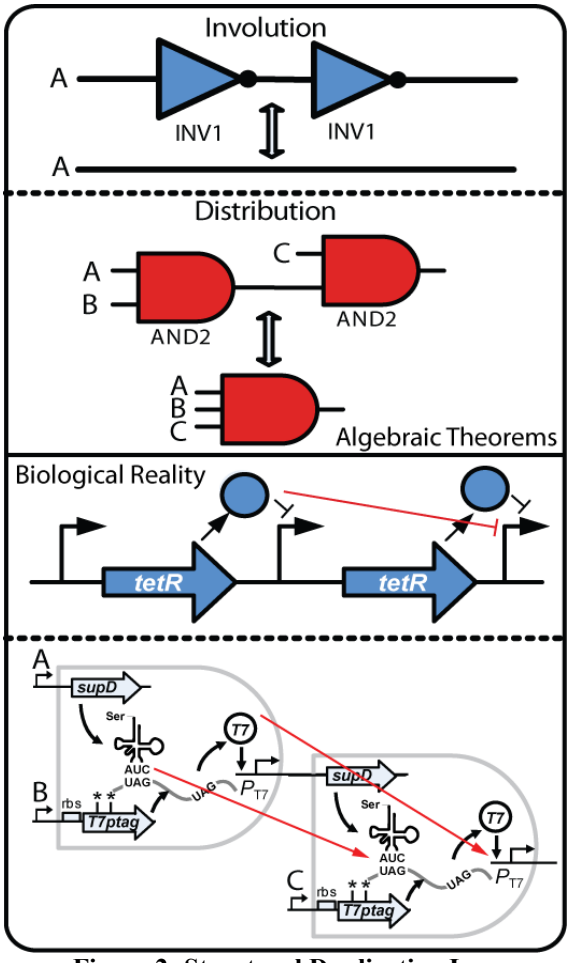

Figure 2: Structural Duplication Issues

currents through a DNA or through a protein or across a membrane, there can be the exogenous addition or production of small molecules, protein-protein interactions, rates of protein or mRNA production, and rates and degrees of protein phosphorylation. In essence, each biochemical component of the cell is an information carrier, and indeed there has been significant diversity in the choice of signal carriers in the field of prototypical logic devices. In a realm in which logic is used fairly simply—say one logic device is used within the entire cell, the concerns about cross reactivity and scalability are irrelevant, but if we are to pursue a VLSI-like scaling, it must be possible to introduce many of these logic devices and connect them both in parallel and in series. The inputs and outputs of scalable logic device therefore must be of the same form, and this requirement eliminates a wide swath of options in biology. Small molecule inputs to riboregulator-based and transcription factor-based genetic circuits have been the bulk of devices made, but the outputs of these devices have all been changes in transcription rates, translation rates, or protein activity levels. The architecture of most of these devices cannot be generalized to a common input/output form, and they are therefore not amenable to the production of scalable combinatorial logic. There are, however, at least two signal carriers that might be amenable to this type of logic: generalized transcriptional signals, and phosphorylation signals.

PoPS (Polymerases Per Second), or a generalized transcription rate, has been proposed by Endy [7] and coworkers as an equivalent of electronic current for biology. In general, most behavior in the cell requires transcription to occur and can therefore be regulated through the manipulation of transcription rates. Moreover, the connectivity of the DNA elements that initiate transcription, called promoters, and the elements that are transcribed, is well known to be modular and spatially dependent: the promoter must be directly 5' to the element that will be transcribed under its control. Common inputs and outputs of generalized transcriptional signals would seem to present a good way of connecting genetic logic devices.

There is, however, one major constraint to using PoPS as a signal carrier: they can only be constructed in series. DNA is one dimensional, and only one dimensional relationships can be implemented in DNA. For example, in Figure 3 three PoPS-based NOT gates are in series. As drawn, the biochemical distinctiveness of each of the three NOT gates is distinguishable since they are based on distinct components. Moreover, their physical connectivity makes this equivalent to having placed 3 NOT gates in series, and one would predict that this circuit is equivalent to a single NOT gate. However, when we try to rearrange things to be in parallel, things are complicated by the absence of a second dimension on the DNA. To make things in parallel, one can rewire the devices using components from within the "black boxed" devices, in this case Plac promoter. In effect one is creating a second signal carrier in this case of a particular soluble biochemical species (the transcription factor LacI). However, in doing so the ability to abstract the primitive logic device was lost.

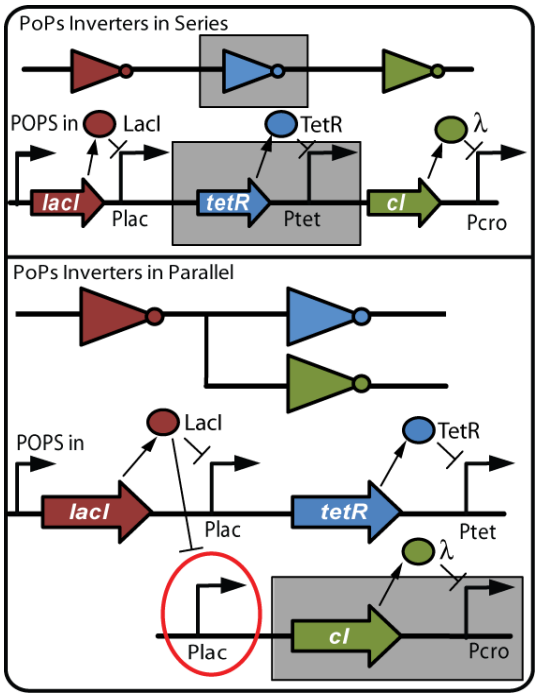

Figure 3: PoPS Inverter Configurations

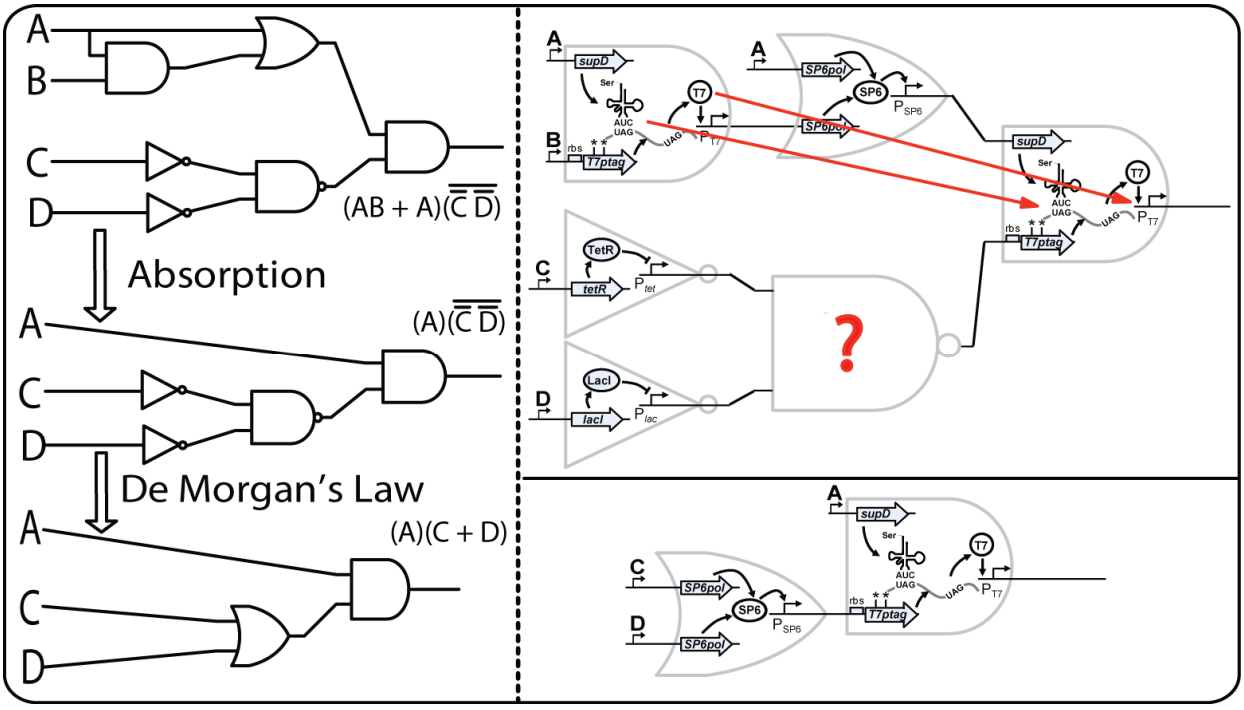

Figure 4: Boolean Algebraic Transformations to Realize Biological Circuitry

Figure 4 illustrates that by applying Boolean algebraic manipulations to the overarching logic of what one wishes to design, certain biological issues can be overcome. On the left portion of the figure three equivalent circuits are shown along with their transformations. On the right are two theoretical biological circuits. Assuming the top circuit cannot be realized due to structural duplication and the lack of a designed NAND gate, one could imagine that the bottom circuit could remove these issues by using two distinct biological structures.

IV. Conclusions and Future Work

A variety of barriers currently prevent us from readily producing scalable logic devices using synthetic biology. However, much work is underway to get beyond these limitations. Efforts have begun to develop scalable RNA logic devices that employ distinctive Watson-Crick base pairing addresses to biochemically distinguish individual devices within a circuit [8]. Similarly, methods of scaffolding protein-protein interactions to direct phosphorylation signals are in active development [9]. Finally, one attractive resolution to the scalability issues of these devices is to consider cells themselves as the primitive logic unit. Here, networks of spatially-addressed cells communicating via cell-cell interactions replace networks of genetic devices communicating through biochemical interactions. In addition to the architectural constraints on the design of scalable logic, there are a number of technical constraints on the physical construction of such systems that must be addressed prior to its successful implementation. These include issues such as toxicity and the dramatically different switching times of biological circuits versus their digital counterparts. Fortunately, much work in synthetic biology is currently focused on developing scalable methods for physical DNA assembly and developing methods for encapsulating and modeling the quantitative behavior of genetic circuits. While synthetic biology is still in its infancy, we remain confident about its potential and the role that traditional VLSI CAD techniques can play in its development.

Acknowledgment

Thanks to the Synthetic Biology Engineering Research Center (SynBERC) for fostering this collaboration.

References

[1] 1. Anderson, J.C., C.A. Voigt, and A.P. Arkin, Environmental signal integration by a modular AND gate. Mol Syst Biol, 2007. 3: p. 133.

[2] 2. Cox, R.S., 3rd, M.G. Surette, and M.B. Elowitz, Programming gene expression with combinatorial promoters. Mol Syst Biol, 2007. 3: p. 145.

[3] 3. Guet, C.C., et al., Combinatorial synthesis of genetic networks. Science, 2002. 296(5572): p. 1466-70.

[4] 4. Ashkenasy, G. and M.R. Ghadiri, Boolean logic functions of a synthetic peptide network. J Am Chem Soc, 2004. 126(36): p. 11140-1.

[5] 5. Rackham, O. and J.W. Chin, Cellular logic with orthogonal ribosomes. J Am Chem Soc, 2005. 127(50): p. 17584-5.

[6] 6. Rinaudo, K., et al., A universal RNAi-based logic evaluator that operates in mammalian cells. Nat Biotechnol, 2007. 25(7): p. 795-801.

[7] 7. Endy, D., Foundations for engineering biology. Nature, 2005. 438(7067): p. 449-53.

[8] 8. Lucks, J.B., et al., Toward scalable parts families for predictable design of biological circuits. Curr Opin Microbiol, 2008.

[9] 9. Dueber, J.E., et al., Rewiring cell signaling: the logic and plasticity of eukaryotic protein circuitry. Curr Opin Struct Biol, 2004. 14(6): p. 690-9.