Math 197: Senior Thesis

Minimal Circuits for Some Classes of Functions

Richard Strong Bowen rbowen@hmc.edu

Harvey Mudd College

Mathematics @ Harvey Mudd College

Introduction

The field of circuit complexity is concerned with finding small circuits to compute various functions. This thesis began by consideration of minimal circuits for *incompletely specified* functions. The goal was to find the (asymptotic growth of) the minimum number of circuit elements needed to compute a boolean function on n variables where only some small number of the required outputs are specified, and where a certain error is allowed. Pippenger (1976) solved this problem for the case that the number of specified outputs is a fixed fraction of the 2^n outputs; Sholomov (1969) solved the same problem for slower-growing numbers of specified outputs but with no allowed error. We showed the combined theorem, solving the problem both with error and slower-growing numbers of specified outputs.

Additionally, we explored several other restrictions on functions and their effects on minimal circuit size.

The problem of partially-specified functions is motivated in part by theoretical biology. Livnat and Pippenger (2008) are interested in the behavior of resource-limited computation devices as a model for organisms' brains.

Circuits

By a *circuit* we mean an acyclic network of circuit elements, each of which computes a function. These are combined to create more complex functions. Each circuit has some n inputs and one output. In this formalism, we can study the number of gates needed to compute various functions

- The *complexity* $L(f)$ of a boolean function f is the minimum number of circuit elements required to make a circuit computing it.
- The complexity $L(S)$ of a set of boolean functions S is the maximum complexity of any function in S:

$$L(S) = \max_{f \in S} L(f).$$

Hence, $L(S)$ gates is sufficient compute any function in S.

How Bounds are Proven

We are interested in the size of minimal circuits for a class of boolean functions as we increase the number of inputs.

Lower Bounds

To prove a lower bound of q, we mean we proving a theorem similar to the following: let p be a property of boolean functions, and P_n be the set of n-input boolean functions having property p. Then $L(P_n) \gtrsim q(n)$.

In general, lower bounds are proven by a counting argument (see for example Wegener (1987)). Two functions can't be computed by the same circuit, so to prove a lower bound we count the number of functions in the class, and the number of circuits with a number n of elements. To get every circuit, we must at least increase n until the latter is greater than the former.

Upper Bounds

To prove an upper bound of q, we mean proving a similar to the following: let p be a property of boolean functions, and P_n be the set of n-input boolean functions having property p. Then $L(P_n) \lesssim q(n)$. Upper bounds are usually proven by explicit construction – that is, by giving a method which will always construct a small enough circuit for any function in P_n. They are typically more difficult to prove than lower bounds.

Results

Our primary result is an upper bound for the case of very incompletely-specified boolean functions. Additionally, we have proven several lower bounds, and are close to a second upper bound.

Very Incompletely Specified Functions

Pippenger (1976) studied incompletely specified functions, that is, a class of functions parameterized by a fraction p indicating how many function values are specified, as well as allowing a fraction E of errors. In this class, $p2^n$ function values are specified; these were referred to as *incompletely specified functions.*. In a similar method to Sholomov (1969), we extend this result to include more slowly-growing numbers of specified values; these we call *very incompletely specified functions.*

The statement of the theorem is: let P_n be the set of n-input functions specified on R_n inputs, and allow a fraction E of errors. Suppose

$$R_n \geq n \log_2^{1+\delta} n$$

for some δ. Then

$$L(P_n) \sim (1 - H(E)) \frac{R_n}{\log R_n}.$$

Here H is the binary entropy function

$$H(p) = -p \log p - (1-p) \log(1-p).$$

We prove the theorem using a very similar process to Sholomov (1969). The lower bound can be proven by a counting argument. The upper bound is proven by constructing a circuit for a given function f. Since f is incompletely specified, we complete it with a function g. The function g may also differ from f in a fraction E of places. g is chosen by a covering lemma to have a minimal entropy; hence it can be described in a short binary string, $\chi(g)$. The length of $\chi(g)$ is short enough to allow it to be in a sense hard-coded into a circuit while using few enough gates to match the lower bound.

Investigating the Factoring Behavior

Sholomov (1969) showed the number of gates required for a function specified in R_n places is $\frac{R_n}{\log R_n}$. Pippenger (1976) showed that the number of gates required for a function when E errors are allowed is $(1 - H(E))\frac{2^n}{n}$. The expression $\frac{2^n}{n}$ is the number of gates required for a general circuit of size n. Comparing these to our derived formula of

$$(1 - H(E)) \frac{R_n}{\log R_n}$$

as well as several other similar bounds in literature, an immediate question of whether this "factoring" behavior persists. We began an investigation by considering a pair of restrictions, choosen because both individually or when applied together, the counting is simple enough to easily find lower bounds. No upper bounds have been completely proven using the principle of local coding, although we have made significant progress on skewedness.

Conclusions and Future Work

The most interesting queston for future work is: when exactly does the factoring behavior occur? The lower bounds derived don't suggest an easy parameterization for factoring in the case we have examined, but much more work is needed to have an exact characterization of this behavior. Additionally, given the kind of modifications to Sholomov's arguments, it seems that a more general theorem could be stated for the upper bound.

References

Livnat, A., and N. Pippenger. 2008. Systematic mistakes are likely in bounded optimal decision-making systems. *Journal of Theoretical Biology* 250(3):410–423.

Pippenger, N. 1976. Information theory and the complexity of Boolean functions. *Mathematical System Theory* 10(1):129–167.

Sholomov, L. A. 1969. On the Realization of Incompletely-Defined Boolean Functions by Circuits of Functional Elements. *Problemy Kibernetiki* 10:215–226. Trans: System Theory Research, 21 (1969) 211-223.

Wegener, I. 1987. *The complexity of Boolean functions*. Teubner.

Acknowledgments

I wish to express my thanks to my advisor, Professor Nick Pippenger;to my second reader, Professor Ran Libeskind-Hadas; to all the math department faculty; and to the math department staff, especially the system administrator, Claire Connelly.

Advisor: Nick Pippenger **Reader:** Ran Libeskind-Hadas