数理解析研究所講究録 1426 巻 2005 年 96-100

An Energy Complexity Measure for Threshold Circuits that is Motivated by Biological Data on Cortical Computations[1]

東北大学大学院・情報科学研究科　内沢　啓 (Kei Uchizawa)
Graduate School of Information Sciences,
Tohoku University

Abstract

We propose new complexity measure, *energy complexity*, for artificial circuits of threshold gates (= McCulloch-Pitz neurons or perceptrons) that measures the average number of intermediate gate outputs "1" over all the inputs. This complexity measure is motivated by the fact in biology such that neurons of biological neural circuits consumes substantially more energy when those output the analogous value "firing" than "non-firing". Hence biological circuits need to save the number of "firing" inrder to use limited energy effectively. We show that a fairly large class of computational problems can be solved by feedfoward threshold circuits of polynomial size and only $O(\log n)$ energy complexity.

1 Introduction

A recent biological study on the energy cost of cortical computation [2] concludes that "The cost of a single spike is high, and this limits, possibly to fewer than 1 %, the number of neurons that can be substantially active concurrently". According to the best currently available estimates the fraction of neurons in cortex that are firing during a typical cortical computation that lasts up to 1 second is about 1 %. In contrast to that, one has to assume that about 50 % of gates in a typical feedforward threshold circuit, also in a circuit that has been

[1]The work was partially supported by the Austrian Science Fund FWF, project # P15386.

optimized with regard to traditional circuit complexity measures, output the value 1 during a computation. Consequences of this sparse firing activity of neurons in cortex have been already addressed in numerous theoretical studies of computational neuroscience, but only in terms of corresponding schemes for sparse representations of information (which are also advantageous for other reasons, see [1] for a recent review). Obviously these studies of biologically realistic sparse data structures in neural circuits need to be complemented by the investigation of algorithms. In this study we introduce new complexity measure, *energy complexity*, that measures the average number of intermediate gate outputs "1" and try to investigate feedfoward threshold circuits from that perspective.

2 Definition of Energy Complexity

We consider circuits consisting of threshold gates (also called linear threshold gates, McCulloch-Pitts neurons, or perceptrons). We set $\mathrm{sign}(z) = 1$ if $z \geq 0$ and $\mathrm{sign}(z) = 0$ if $z < 0$. A threshold gate g (with weights $w_1, \ldots, w_n \in \mathbb{R}$ and threshold $t \in \mathbb{R}$) gives as output for any input $X = (x_1, \ldots, x_n) \in \{0,1\}^n$

$$g(X) = \mathrm{sign}(\sum_{i=1}^{n} w_i x_i - t) = \begin{cases} 1, & \text{if } \sum_{i=1}^{n} w_i x_i \geq t \\ 0, & \text{otherwise} \end{cases} .$$

For a threshold gate g_i within a feedforward circuit C that receives $X = (x_1, \ldots, x_n)$ as *circuit input* we will write $g_i(X)$ for the output that gate g_i gives for this circuit input X (although the actual input to gate g_i during this computation will in general consist of just some bits from X and/or outputs of other gates in the circuit C). With this convention we define the *energy complexity* of a circuit C consisting of threshold gates $g_1, \ldots, g_m$ as

$$EC(C) := \frac{1}{2^n} \sum_{X \in \{0,1\}^n} \sum_{i=1}^{m} g_i(X) .$$

We will write $G(X)$ for *activation pattern* $(g_1(X), \ldots, g_m(X))$ of the circuit C for input $X \in \{0,1\}^n$. Furthermore we classify the activation patterns into two sets by the value of circuit output $C(X) \in \{0,1\}$ and define:

$$AP_j(C) := \{G(X) : X \in \{0,1\}^n \text{ and } C(X) = j\}$$

for $j \in \{0,1\}$.

3 Result

Obviously the size of the total activation patterns $AP(C) = AP_0(C) \cup AP_1(C)$ of a circuit C is somewhat related to its energy complexity $EC(C)$, since an upper bound on $EC(C)$ and on the total number of gates in the circuit C entails an upper bound for $AP(C)$. This results from the simple combinatorial fact that a bound on the number of 1's in a bit string of length m gives rise to an upper bound on the total number of such bit strings of length m. For example, if one assumes that at most $k{\cdot}\log n$ many 1's occur in any computation of circuit C for some input $X \in \{0,1\}^n$ then $\#AP(C) \le \sum_{j=0}^{k} \binom{m}{j}$. So for reasonably small k and m we have a nontrivial upper bound on $\#AP(C)$. On the other hand there exists a extra case such that $\#AP(C) = 1$ and $EC(C) = m$. But theorem 1 below state that it's possible to convert C to C' with polynomial blow-up in size and small $EC(C')$, precisely:

Theorem 1: Assume that C is a threshold circuit computing a Boolean function f: $\{0,1\}^n \to \{0,1\}$ with m gates. Then one can construct a depth $m+1$ threshold circuit C' computing f with at most $mp+1$ gates and $EC(C') \le \log p + 2$ where $p = \min(\#AP_0(C), \#AP_1(C))$.

where $\#S$ for a set S is the cardinality of S. Theorem 1 gives the following result,

Corollary 1: Assume that C is a threshold circuit computing a Boolean function f: $\{0,1\}^n \to \{0,1\}$ with $poly(n)$ gates and $\min(\#AP_0(C), \#AP_1(C)) = poly(n)$. Then one can construct threshold circuit C' computing f with $poly(n)$ gates and $EC(C') = O(\log n)$.

Many threshold circuits for computing specific functions that have been discussed in the literature (for example threshold circuits of depth 2 that compute symmetric functions in linear size) produce just $O(n)$ many activation patterns, but also on average $\Omega(n)$ 1's. Corollary 1 state that it's possible to compute such Boolean functions using threshold circuits with only $O(\log n)$ energy complexity. On the other hand the depth of the circuit C' after this conversion is increased to the size of original circuit C even if the depth of C is a constant.

We don't know well how strong the restrictions in corollary 1 are, since the activation pattern is new concept for threshold circuits. But the corollary 2 from theorem 2 partially answers this question.

Theorem 2: Let C be a threshold circuit computing a Boolean function f:$\{0,1\}^n \to \{0,1\}$ with m gates where the absolute value of a weight are between 1 and w. Then one can construct a depth 3 threshold circuit C' computing f that consists of $n^{O(\log(m+n)w \log p \log\log p)}$ gates $p = \min(\#AP_0(C), \#AP_1(C))$.

Theorem 2 can be proved using similar conversion technique used in [3]. Theorem 2 gives the following corollary,

Corollary 2 Let C be a threshold circuit computing a Boolean function f:$\{0,1\}^n \to \{0,1\}$ with $poly(n)$ gates where $min(\#AP_0(C), \#AP_1(C)) = poly(n)$ and weights are between 1 and $O(n^{poly(\log n)})$. Then one can construct a depth 3 threshold circuit computing f with $O(n^{poly(\log n)})$ gates.

Corollary 2 is interesting also from the viewpoint of circuit complexity theory, because we don't know a Boolean function that cannot be computed by a depth 3 threshold circuit while [4] showed there exists a Boolean function that cannot be computed by a depth 2 threshold circuit. From this pesperctive corollary 2 seems significant.

4 Future Work

Since the energy complexity is new concept for threshold circuits there exist many directions. As we said in Section 1 it's attractive to study the sparse firing activity in biological circuit using the complexity measures introduced here. Another direction is to investigate upper and lower bounds of the energy complexity (and also the activation pattern complexity) for a specific Boolean function. It is also important to investigate a relationships between these new measures and conventional complexity measures like size or depth.

5 Acknowledgements

We would like to thank Rodney Douglas and Robert Legenstein for helpful discussions, Kazuyuki Amano and Eiji Takimoto for his advice, and Akira Maruoka for making this collaboration possible.

References

[1] Földiak, P.(2003). Sparse cording in the primate cortex., *The Handbook of Brain Theory and Neural Networks*, 1064-1068. MIT Press.

[2] Lennie, P.(2003). The cost of cortical computation. *Current Biology*, 13:493-497.

[3] Klivans, A, R. and Servedio, R, A.(2004). Learning DNF in Time $2^{\tilde{O}(n^{1/3})}$. *JCSS* 68(2), 303-318.

[4] Hajnal A., Maass W., Pudlak P., Szegedy M, and Turan G.(1993) Threshold circuits of bounded depth. *JCSS* 46(2), 129-154.