氏名（本籍地）	内沢　啓（うちざわ　けい）
学位の種類	博　士（情報科学）
学位記番号	情　博　第　392　号
学位授与年月日	平成20年3月25日
学位授与の要件	学位規則第4条第1項該当
研究科、専攻	東北大学大学院情報科学研究科（博士課程）情報基礎科学専攻
学位論文題目	Energy Complexity of Threshold Circuits （しきい値回路のエネルギー複雑度に関する研究）
論文審査委員	（主査）東北大学教授　西関　隆夫 東北大学教授　篠原　歩　　東北大学教授　徳山　豪 東北大学准教授　瀧本　英二

論文内容の要旨

1. Introduction

Biological neurons in the brain communicate with each other via electrical signals, called "spikes," in order to perform various tasks, e.g., cognition, learning, memory, etc. A neuron emitting the spikes is said to be "firing." A recent biological study reports the following interesting fact about the energy consumption of the neural firing in the brain: the energy cost of a single spike is very high, and so the energy consumed for the spikes is a large fraction of the total energy supplied to the brain. Thus information processing in the brain results from a low firing rate of neurons. Such low activity of neural circuits is sometimes referred to as *sparse activity*. Some biological researches insist that the sparse activity may offer several advantages for the brain (besides conserving energy), and constitute a general principle of information processing in the brain. Now the following natural questions arise:

q1. How do neural circuits save the number of firing neurons when performing certain computational tasks?

q2. To what extent is the computational power of neural circuits restricted if the number of firing neurons is bounded?

In the thesis, we consider the two questions above from the viewpoint of *circuit complexity*, and try to answer them. Circuit complexity is a discipline of computational complexity theory which studies the inherent difficulty of computational problems. In computational complexity theory, we consider a particular computational model, in which we measure the difficulty of computational problems in terms of the amount of computational resources needed to solve them. In circuit complexity, the models are feedforward circuits of gates, and the resources are typically the size and the depth of circuits, where the *size* of a circuit is the number of gates contained in the circuit and its *depth* is the length of a longest path from an input gate to the output gate. The amount of resources needed to solve problems depends on what class of functions each gate can perform. Therefore, in order to discuss theoretically our two questions, we first determine the class of gate functions suitable for describing neural computations.

We employ, in the thesis, the class of linear threshold functions as the gate functions, since a linear threshold function approximately represents the input-output characteristics of a biological neuron. A gate computing a linear threshold function is called a *threshold gate*, and a circuit consisting of threshold gates is called a *threshold circuit*. A threshold gate has two output states, 1 and 0, and we consider an output "1" to be neural firing. Among many types of computational models of neural circuits, a threshold gate is, historically and even currently, one of the common abstract computational models of a biological neuron, and a threshold circuit is that of a neural circuit. There are a number of researches on the computational power of threshold circuits, and those research present many results.

However, the number of firing gates during information processing gets less attention in previous research of threshold circuits. In fact, threshold circuits constructed in the previous research are highly likely to have the property that a large portion of gates in the circuits fire during computation. This is because the main concern of circuit complexity has been to minimize the particular computational resources, i.e., size and depth, in constructing circuits that compute Boolean functions. This discrepancy stems from a fundamental difference in the energy cost of computations in the brain and computations in electronic circuits. Common abstract measures for the energy consumption of electronic circuits treat the two output states 1 and 0 of a gate symmetrically, and focus instead on the required number of switching's between these two states. Usually, a measure for energy consumption of optical computing was proposed, but it also treats the two output states symmetrically. Thus, people have never cared about the number of the gates that output "1," since we tend to think of circuits as being implemented electronically or optically.

Therefore, we initiate our study by introducing a new complexity measure, *energy complexity*, whose minimization yields computations with sparse activity. Roughly, energy complexity of a threshold circuit is defined as the number of firing gates in the circuit during computation. Now we rephrase our two questions **Q1** and **Q2** in terms of threshold circuits and energy complexity:

Q1. How do we design threshold circuits of small energy complexity for computing certain Boolean functions?

Q2. To what extent is the computational power of threshold circuits is restricted if the energy complexity of the circuits is bounded?

In the thesis, we partially answer these questions.

2. Energy-Conscious Model of Threshold Circuits

We consider a threshold gate having an arbitrary number m of fan-in's as a model of a neuron. For every input $z = (z_1, z_2, \ldots, z_m) \in \{0, 1\}^m$, a *threshold gate* g (with *weights* $w_1, w_2, \ldots, w_m$ and a *threshold* t) computes a linear threshold function $g(z)$ such that $g(z)=1$ if the sum of $w_i z_i$ for all i is greater than 0, otherwise $g(z)=0$. We assume that the weights and threshold of every threshold gate are integers. We consider the output 1 as the firing.

A *threshold circuit* C with n input variables is represented by a directed acyclic graph; the graph has exactly n nodes of in-degree 0, each associated with an input variable and called an *input node*; each of the other nodes represents a threshold gate. For an assignment $x \in \{0, 1\}^n$ to the n input variables, the output of all gates in C are computed in topological order of the nodes in the directed acyclic graph. For a gate g in C, we denote by $g[x]$ the output of g for an input x to circuit C. Since we consider only a threshold circuit that computes a Boolean function, one may assume without loss of generality that the circuit has exactly one gate of out-degree 0, called the *top gate*. We denote by $C(x)$ the output of the top gate of C for x. We say that a threshold circuit C *computes* a Boolean function $f: \{0, 1\}^n \rightarrow \{0, 1\}$ if $C(x) = f(x)$ for every input $x \in \{0, 1\}^n$.

The *size* of a threshold circuit C is the number of gates in C. The *depth* of C is the length of a longest path to the top gate of C.

Now we introduce two measures for the energy consumption of threshold circuits. One is the *maximum* energy complexity, and the other is the *expected* energy complexity. We give precise definitions of the two measures below.

We first define the maximum energy complexity.

Definition 1. *Let C be a threshold circuit of n input variables and s threshold gates $g_1, g_2, \dots g_s$ for some numbers n and s. The maximum energy complexity of C, denoted by $EC_{\max}(C)$, is defined to be the maximum number of gates fired by inputs x where the maximum is taken over all input assignments.*

We then define the expected energy complexity as follows.

Definition 2. *Let C be a threshold circuit of n input variables and s threshold gates $g_1, g_2, \dots, g_s$ for some numbers n and s. Let Q be a distribution over the input assignments for the circuit C. The expected energy complexity of C with respect to Q, denoted by $EC_Q(C)$, is defined to be the expected number of gates fired by inputs x where x is chosen from $\{0, 1\}^n$ according to the probability distribution Q.*

3 Constructions of Circuits with Small Energy Complexity

In the chapter, we present results that partially answer Question **Q1**. More specifically, we give constructions of threshold circuits from linear decision trees so that the energy complexity of the resulting circuit is significantly small if the given trees are reasonably small, where *linear decision trees* are binary decision trees such that the classification rule at each internal node is performed by a threshold function.

We first consider the case of the maximum energy complexity. The following theorem shows how small the energy complexity of the resulting circuit is.

Theorem 1. *Let f be a Boolean function computable by a linear decision tree of L leaves. Then f is also computable by a threshold circuit C of size L such that $EC_{\max}(C) \leq \log L + 1$.*

Note that the resulting threshold circuit is of a polynomial-size and energy complexity $O(\log n)$ if the target Boolean function is computable by a linear decision tree of polynomial leaves. Thus, Theorem 1 implies that threshold circuits have considerably large computational power even if we restrict the energy complexity to the logarithm of its size.

Furthermore, we develop a more refined analysis with respect to the expected energy complexity, and give a similar theorem on the construction of threshold circuits of small expected energy complexity. In this case, the expected energy complexity of the resulting circuit with respect to Q is bounded above by a cost of a linear decision tree, where the *cost* is defined to be the expected number of 1s that linear threshold functions outputs at nodes in the tree, where the expectation is taken over the probability distribution Q. More specifically, we prove the following theorem:

Theorem 2. *Let f be a Boolean function computable by a linear decision tree of L leaves, and Q be a probability distribution for input assignments. Then f is also computable by a threshold circuit C of size L such that* $EC_Q(C)$ *is bounded above by the cost of T over Q.*

The constructions of threshold circuits are based on the same technique: We somehow embed the structure of a linear decision tree in threshold circuits so that only the gates on a path are allowed to be fired for any input assignment. This technique provides a general methodology of designing systems of energy-efficient computation.

4 Trade-off among the Three Complexities

Our second result is a partial answer for Question **Q2**, and is expressed as a trade-off among the three complexities, size, depth, and maximum energy complexity, in the sense that the three resources cannot be simultaneously small. To derive the trade-off, we employ a *communication complexity* argument.

Communication complexity is a complexity measure for Boolean functions, and has been used as a tool for investigating various notions of complexity appearing in a variety of areas. In particular, a number of arguments through the analysis of communication complexity have been developed to derive lower bounds on the size of threshold circuits. We find out a similarity in these arguments and establish a general scheme with communication complexity arguments which gives a unified view of the previously developed ones, although the scheme may be folklore in research area of the circuit complexity.

In line with the general scheme, we derive a trade-off in the form of an upper bound on the communication complexity of a Boolean function computed by a threshold circuit. Our bound is expressed in terms of the three complexities of the circuit and is monotonically increasing with respect to each of them. On the other hand, the communication complexity is determined solely by the Boolean function and is independent of the circuits that compute the function. Therefore, the bound on the communication complexity implies that any of these complexity measures cannot take a smaller value, unless another measure takes a larger value. More specifically, we have the following theorem.

Theorem 3. *Let f be Boolean function of n variables such that the communication complexity of f is* $O(n)$*. Then, any threshold circuit computing f satisfies* $(e + d)^d \log s = \Omega(n)$*, where s, d, and e are size, depth, and energy complexity of the circuit, respectively.*

In fact, since almost all Boolean functions have a linear communication complexity, we can conclude that a large class of Boolean functions has the trade-off given in Theorem 3.

5 Discussions

Recall the theorems in Chapter 3. Although we do not describe constructions giving the theorems in the paper, every of the resulting circuits constructed from the theorems actually has two properties: a tree-like structure and a large depth.

By the tree-like structure, the circuit explicitly has a property that, after each evaluation of outputs of threshold gates for an input, the gates "inhibit" gates that do not influence on the future processing to have less energy complexity. Thus, we observe that the tree-like structure is one of effective strategies to decrease the firing rate. In other words, neural circuits in the real brain may propagate information what neurons should have inhabitation in order to achieve sparse activity, and consequently, it could be a factor for increasing the depth.

While our constructions yield threshold circuits of small energy complexity, the circuits tend to have a large depth which seems a week point of our construction. However, it may be inevitable for threshold circuits of small energy complexity to have a large depth, because the trade-off given in Theorem 3 implies that the depth seems to be more negatively correlated to the energy complexity than its size.

In order to avoid increasing depth, but to obtain small energy complexity, one can construct threshold circuits of small energy complexity by taking another strategy such that a significantly large number of gates perform a target task. However, if we consider a case of real neural circuits, it may be reasonable to assume that a bounded number of neurons in the brain are available for a neural circuit to perform a specific task. In that case, the trade-off directly implies a strong relation between the depth and the firing rate, and that neural circuits may need large depth to achieve low firing rate.

論 文 審 査 結 果 の 要 旨

脳はニューロンの疎な発火活動によりエネルギー消費効率のよい情報処理を行っていることが，最近の脳科学研究から明らかになりつつある。脳における情報処理のしくみを理論的に解明するために，しきい値回路が脳の数学モデルとして古くから用いられているが，これまで回路のエネルギー消費効率が考慮されることはなかった。そこで著者は，しきい値回路モデルに対し，エネルギー複雑度という新しい尺度を導入し，エネルギー複雑度を制限したしきい値回路について，一般的な設計手法を与えるとともに，その計算能力に一定の限界があることを示した。本論文はこれらの成果をとりまとめたもので，全編５章からなる。

第１章は序論である。

第２章では準備として，しきい値回路モデルの定義を与えるとともに，しきい値回路において１を出力する素子の個数の最大値を，その回路のエネルギー複雑度と定義している。エネルギー複雑度は脳のエネルギー消費量をモデル化したものであり，エネルギー複雑度の導入は計算理論と脳科学の双方の分野においてきわめて有用である。

第３章では，エネルギー複雑度の小さいしきい値回路の設計手法を２つ与えている。１つ目では，ブール関数が線形決定木と呼ばれる表現形式で与えられたとき，その関数を実現するしきい値回路で，エネルギー複雑度が素子数の対数程度であるものを構成している。これは，エネルギー複雑度が非常に小さいしきい値回路でも，十分高い計算能力を有していることを意味しており，重要な成果である。２つ目では，任意のしきい値回路が与えられたとき，その回路と同じブール関数を実現するエネルギー複雑度の小さいしきい値回路を構成している。この結果は，エネルギー消費効率の優れたしきい値回路の設計基盤を与えるだけでなく，脳の計算原理の一端を解明する手がかりを与える可能性があり，高く評価できる。

第４章では，しきい値回路の素子数，層数，およびエネルギー複雑度の３つのパラメータの間にはトレードオフがあり，同じブール関数を実現するしきい値回路で，これらのパラメータがいずれもある限界値以下となるような回路は存在しないことを示している。この結果は，脳が優れたエネルギー消費効率を得る代償として，層数，すなわち，計算時間を犠牲にしている可能性を示唆しており，きわめて興味深い。また，この結果から，層数とエネルギー複雑度が制限されると，内積関数など特定のブール関数を実現するしきい値回路の素子数は指数的になることを導いている。これは優れた結果である。さらに，この結果を導出するときに用いた通信複雑度の理論は，様々な計算量の下界を証明するための新しい枠組みを与えるものとして，きわめて有用である。

第５章は結言である。

以上要するに本論文は，エネルギー消費効率という視点を計算理論の分野に初めて取り入れたきわめて独創性の高いものであり，エネルギー消費効率のよいしきい値回路の計算原理について新しい知見を与えるもので，情報科学，特に計算理論の発展に寄与するところが少なくない。

よって，本論文は博士（情報科学）の学位論文として合格と認める。