Molecular Neuron Realization

Piotr Wąsiewicz, Andrzej Dydynski
Grzegorz Tomczuk, Jan J. Mulawka
The Institute of Electronic Systems

Andrzej Płucienniczak
The Institute of Biotech. & Antibiotics

Abstract

Universal Turing machine is a notional computing machine that stimulated work leading to the development of modern computers. The Turing machine operates on finite sequences of symbols by scanning a data type. The striking analogy to information-encoding biochemical reactions on information-carrying molecules was inspired to apply this methodology in neural networks. The essential feature of such approach is hybridization of pairs of complementary DNA strings and possibility to represent highly parallel selective operations, which can enable creating alternative, neural architectures. We describe our original model of molecular neuron based on DNA computing paradigm. During computation appropriate molecules are chosen, each specifying one of the finite number of initial states or processing elements.

Keywords: **DNA computing, molecular neuron, genetic engineering, neural network, molecular computation**

1 Introduction

MODERN computer architectures are based on Turing machine. Such solution has one processor. It operates in sequential order executing in one moment one operation after another operation. The list of operations is called an algorithm. However, inspired by biology methodologies like neural networks, evolutionary programming and so on consist of many processing units. Their performance cannot be described in the form of the sequential algorithm, more precisely in the form of the operation single list suitable for execution on one processor. Moreover, the NP-difficult problems can be solved in polynomial time only by non-deterministic machine utilizing massive parallelism. Choosing in one moment among millions of solutions ought to be non-deterministic. It has been noticed that in future computing solutions the need of creating new alternative distributed computer architectures is emerging.

A sequence of operations on DNA executed in parallel on DNA strings is called an algorithm. But in the typical DNA computing algorithm this sequence is determined by a model of DNA strings similar to the soft hardware specialized architecture driven here by heating, cooling and connected with them operations on DNA. Together the operation sequence and the model make computation possible.

P. Wąsiewicz, A. Dydyński, G. Tomczuk and J.J. Mulawka are from the Institute of Electronic Systems, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland, {pwasiewi,adydynsk}@elka.pw.edu.pl ; A. Płucienniczak is from the Institute of Biotechnology and Antibiotics, Staroscinska 5, 02-516 Warsaw, Poland

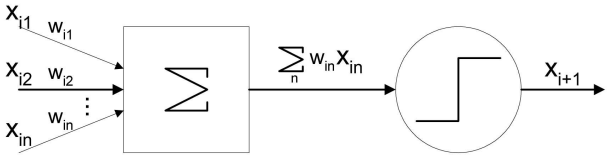

Fig. 1. The McCulloch-Pitts neuron model.

Mills has reported [2] first approach to a neural net representation by using DNA. He explored the possibility that networks in which the usual axons and neurons are replaced by the diffusion and molecular recognition of DNA can make practical use of massive parallelism associated with simultaneous reactions up to 10^{24} molecules. He considered a DNA embodiment of Hopfield neural network or associative memory and designed special form of single-stranded DNA oligomers, called vectors, in order to implement such a neural network. The method of adding such vectors and how to find inner and outer products of two vectors, product of matrix and a vector and product of matrices were considered.

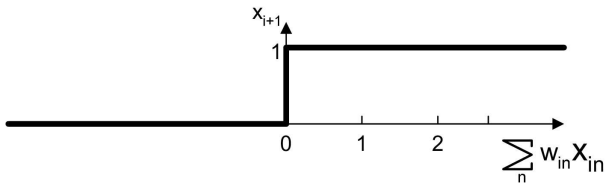

Fig. 2. The McCulloch-Pitts neuron threshold function g.

A Hopfield neural network [3] can be realized by implementing memorization and recall. In an experiment data were presented in the form of images of m pixels. The images were flashed on a micro-array, whose pixels matched the image pixels. The DNA data strands were attached by hybridization to complementary anchored strands. The image data was formed into outer products and the sum of all the outer products becomes the memory matrix. In the experiment roughly two iterations were sufficient to force the mixture into a steady state answer to the query.

In this contribution we consider completely another method to model a neural network. The introduced

representation requires all the data to be discrete. This requirement results in the way the information is represented. In all DNA-based methods chemical operations are performed on strings of DNA nucleotides, which code information. In our concept we need to design a subset of all single-stranded DNA strings adequate to the particular problem and associate it with the input data. Some strings that build neuron units are also identifying data. An essential feature of this approach is that network coefficients and input, output signal values are remembered in one DNA string called a detection string.

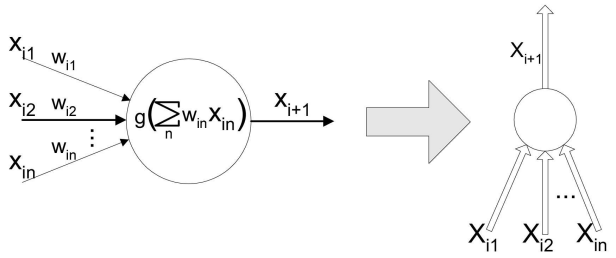

Fig. 3. Creation of the McCulloch-Pitts neuron scheme.

2 The McCulloch-Pitts Neuron Model

McCulloch and Pitts described a simple neuron model as a two-value treshold element. The model consists of two components: neuron and synaptic links. The state of the output signal of a neuron is determined by the linear sum of weighted input signals x_{in}. The output signal x_{i+1} of a neuron is 1 if the sum equals or exceeds the threshold value; otherwise it is 0.

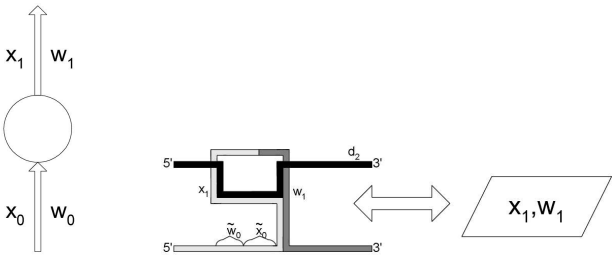

Fig. 4. One cell neuron made of DNA. On the left its traditional symbol. On the right its new schematic symbol.

The McCulloch-Pitts neuron model is shown in Fig. 1 and the threshold with the value equal to 0 is shown in Fig. 2, where $x_{i+1} = g(\sum_n x_{in} \times w_{in})$ The traditional scheme of such the neuron has been created during two steps in Fig. 3. The depicted node contains the threshold function and the adder.

Although its simplicity the McCulloch-Pitts model is a powerful computing tool. McCulloch and Pitts showed that the network formed from such elements is equivalent to the universal computing machine.

3 Realization of a Molecular Neuron by DNA Strings

We consider a class of discrete neural networks, where input and output signals as well as weight coefficients are described by discrete values. In our approach DNA based neuron model is reported in which a standard neuron is formed from oligonucleotide strings resulting in a model that is comparable to that of McCulloch-Pitts one.

TABLE I
The molecular neuron strings with their nucleotide sequences

Strings	Sequence optimization results
X1 47	CACAAGTACACGACATAATCGGAAGCTTC-GCCAGTCTGACGCCGGAT
W1 33	CCGCATCTGCGAATTCGCTGTCGTGTACT-TGTG
D2 6o	GACTAGACCTGGTTCACTGGCGAAGCTTC-CGATTAGCGAATTCGCGTAATGTGTTGGA-GC
Y1 50	CTGATCTGGACCAAGGACTGCGGCCTAGG-CGTAGACATTACACAACCTCG

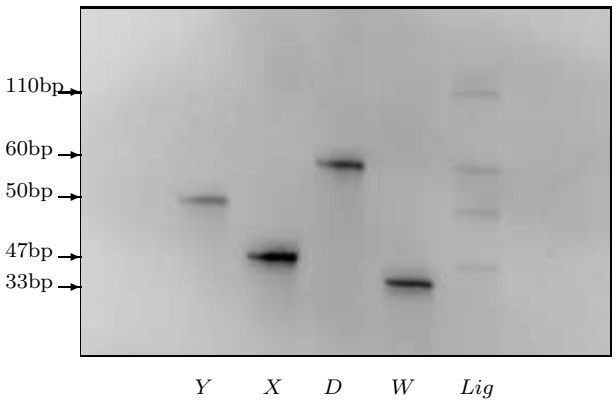

Fig. 5. DNA electrophoretogram of the first experiment: in the *Lig* lane - the evidence of success of our approach, the new strand (110bp) has been created.

The resulting universal cell neuron, which can represent neurons in input, hidden, and output layers, but here with one input and one output signal was introduced in the middle of Fig. 4. Its traditional diagram was introduced on the left side of Fig. 4. On the right side of Fig. 4 the cell neuron was transformed to its new schematic symbol simplifying molecular network descriptions.

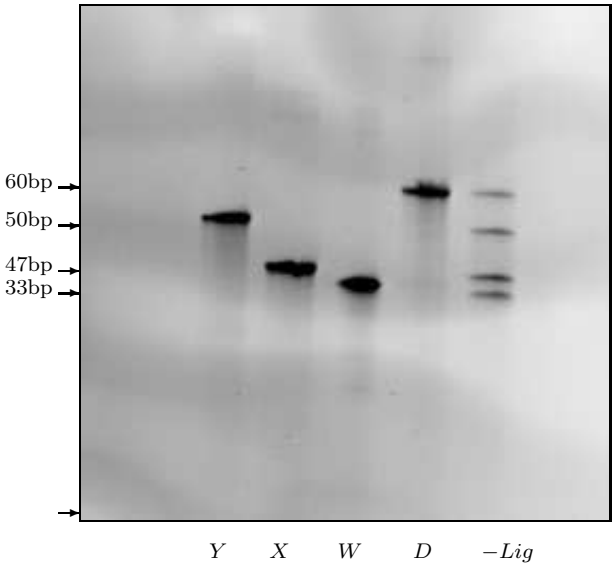

Fig. 6. Experiment without ligation to compare: in the $-Lig$ lane strands on the same level as corresponding Y, X, W, D strands show that there was no ligation.

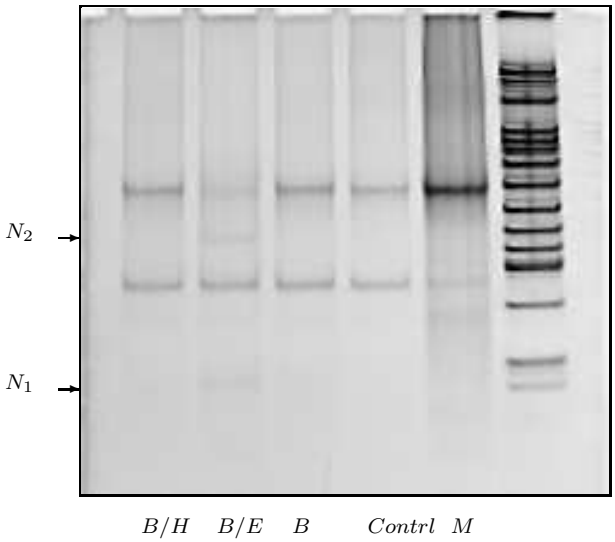

Fig. 7. Digestion results: B/H-digestion by BamHI, B/E-digestion by EcoRI and BamHI, B-Bam only digestion, $Contrl$-control, N_1,N_2-new structures created as products of digetion by both $EcoRI$ and $BamHI$.

We have succeed in implementing a neuron cell, which could be used in building larger and more sophisticated structures like an adaline neural network.

4 Experiment Results and Summary

Concentration of all oligonucleotides was equal to 10 $pM/\mu l$. Before further treatment 150 pM of oligonucleotide W was phosphorylated by $T4$ polynucleotide kinase [1]. Briefly, 150 pM of oligonucleotide W dissolved in 15 μl of TE, pH 8.0 was mixed with 5 μl of 10× $T4$ polynucleotide kinase buffer ($Amersham$), 6 μl 10 mM ATP, 22 μl H_2O and 1 μl (10 units) $T4$ polynucletide kinase ($Amersham$). The reaction was carried out overnight at room temperature. The enzyme was inactivated by heating at 65 ^oC for 10 min. Half of the reaction mixture was mixed with 75 pM of the oligonucleotide X. Separately, equimolare mixture of D and Y oligonucleotides was prepared. 75 pM of each oligonucleotide was present in 30 μl sample containing also buffer for $T4$ DNA ligase ($Amersham$). Both solutions, 75 pM of (W^*+X) and 75 pM $(D+Y)$ oligonucleotides were used to prepare final mixture in 65 μl, containing also $T4$ DNA ligase buffer. The final mixture was heated for 10 min at 37 ^oC and after that 1 μl of $T4$ DNA ligase was added (10 units, USB). The ligation reaction was performed overnight at 16 ^oC.

The results of the ligation reaction were verified during electrophoresis on the 10% polyacrylamid, 7 M urea gel, acrylamid:bisacrylamide = 1:19 (Fig. 5). The results of this experiment clearly show that W and X oligonucleotides are joined together by $T4$ DNA ligase, because after ligation reaction new longer band appears and band corresponding to the W oligonucleotide disappears. In contrary without ligation all strings remain at the same positions (the same lengths) as is shown in Fig. 6.

The ligation mixture (Fig. 5) was also digested with $EcoRI$ and $BamHI$ restriction nucleases (Fig. 7). Part of the mixture was digested with both enzymes e.g. with $BamHI$ where X and W meet (the beginning of W and the last three bp of X) and with $EcoRI$ on the corner of W and D.

The results shown in Fig. 5 and Fig. 7 proof that predicted structure (Fig. 3 and Table. 3) really exists in solution.

References

[1] T. Maniatis, E.F. Fritsch, and J. Sambrook, Molecular cloning: a laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1982).

[2] A.P. Mills, B. Yurke Jr., P.M. Platzman, Article for analog vector algebra computation, *BioSystems* **52** (1999) 175-180.

[3] A.P. Mills, Jr., M. Turberfield, A.J. Turberfield, B. Yurke, P.M. Platzman, Experimental aspects of DNA neural network computation, *Soft Computing, Springler-Verlag* **5** (2001) 10-18.