

CMOS Analog/Digital Circuit of the Hysteresis McCulloch-Pitts Neuron for Ramsey Numbers

YONG BEOM CHO

J.W. Harley, Inc., Twinsburg, OH 44087

KAZUHIRO TSUCHIYA

Fuji Electric Corp. Ltd., Tokyo, Japan 100

YOSHIYASU TAKEFUJI

Department of Electrical Engineering and Applied Physics, Case Western Reserve University, Cleveland, OH 44106

Abstract. A parallel algorithm for finding Ramsey numbers is presented where analog/digital CMOS circuits for the hysteresis McCulloch-Pitts binary neuron are described. The hysteresis McCulloch-Pitts binary neuron model is used in order to suppress the oscillatory behaviors of neural dynamics so that the convergence time is shortened. The proposed algorithm using the hysteresis McCulloch-Pitts binary neuron found five Ramsey numbers. The analog CMOS sigmoid circuit with variable gain controls has been fabricated and tested using the SAC data acquisition board interfaced with a TMS 32010 processor. Hysteresis can be implemented by the positive feedback in the fabricated CMOS analog circuit.

1. Introduction

Recently the silicon implementation of neural networks has received considerable attentions from many VLSI researchers [1], [2]. In this paper analog/digital circuits of the hysteresis McCulloch-Pitts neuron are shown. Through solving NP-complete problems, we have been investigating the feasibility of silicon neural network.

The neural network is composed for neurons and synaptic links. In this paper analog/digital circuits of the hysteresis McCulloch-Pitts neuron are presented which is used to solve Ramsey number problems. The processing element can be modeled using nonlinear transfer function such as a sigmoid-shape transfer function, hard-limit binary function, or hysteresis binary function as shown in figure 1. The processing elements are interconnected by synaptic links.

A silicon neural network can be implemented using analog, digital, or mixed analog/digital circuits. The advantages of an analog circuit are simple basic blocks and smaller area. However, the noise immunity is worse than digital circuits. On the other hand, the digital circuits are straightforward for design and their testability is better than analog circuits. Also the interface to the digital coprocessor is simpler. However, the overall area of digital circuits is larger.

2. Silicon Neuron Models

2.1. Sigmoid Neuron Model

Many optimization problems have been solved by artificial neural networks [3-5]. Some of them used the differentiable and continuous neuron model, sigmoid function, where the input/output relationship is given by

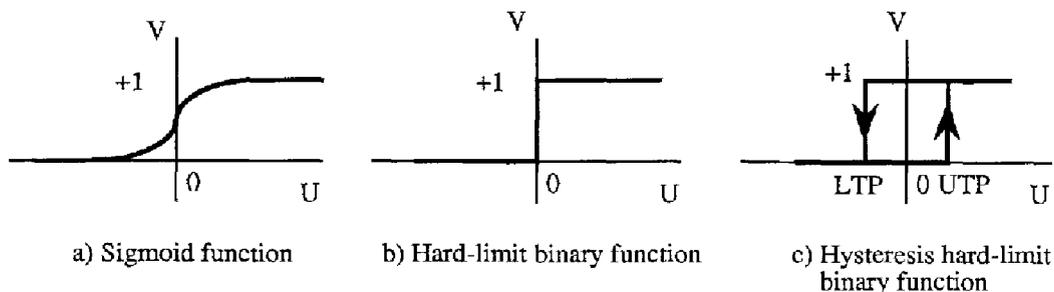


Fig. 1. Nonlinear transfer functions.

$$V_i = g(U_i) = \frac{1}{2} [\tanh(\lambda U_i) + 1] \quad (1)$$

where V_i and U_i are output and the input of the i th neuron, and λ is the gain factor. If λ is infinitely high, the input-output function will become the three-state step function:

$$V_i \text{ is } 1 \text{ if } U_i > 0, \frac{1}{2} \text{ if } U_i = 0,$$

and 0 otherwise.

Figure 2 shows the detailed CMOS circuit diagram of the sigmoid neuron with variable gains. The sigmoid circuit is composed of two subcomponents: an unbuffered comparator with a positive feedback loop and double-cascaded inverters with negative feedback loops [6]. The comparator is provided by transistors $M1$ – $M6$ for the sigmoid function, and the positive and the negative feedback loops in the inverters are provided by transistors $M7$ – $M11$ for gain controls in the sigmoid function.

Figure 3 shows the layout of the fabricated sigmoid neuron circuit using the CMOS technology. It has 10

transistors and 4 resistors which can control the gain of the sigmoid function. The current size is $200 \mu\text{m} \times 150 \mu\text{m}$ based on $3\text{-}\mu\text{m}$ CMOS technology. To evaluate the performance of the proposed circuit, we tested a prototype chip using the SAC data acquisition board interfaced with a TMS 32010 processor. Figure 4 shows the experiments of I/O response on the test chip. The four conditions for r (100, 250, 500, and 1000 where $r = R_{10}/R_9 = R_8/R_7$), are used in the chip. In order to increase the gain of the sigmoid function, the gains of the inverters, either R_{10}/R_9 or R_8/R_7 , or both, should be increased. As the gains of the inverters increase, the I/O sigmoid function curve approaches a high-gain limit function. As the gains are lowered, it approaches a gently-sloped sigmoid curve. In the lowest gains, the output approaches a constant zero voltage, regardless of inputs. In order to reduce dc offset, it was necessary to adjust the V_{SS} to -3.6 V .

2.2. The Hysteresis McCulloch-Pitts Neuron Model

The hysteresis property suppresses the oscillatory behavior of neural dynamics and consequently it shortens

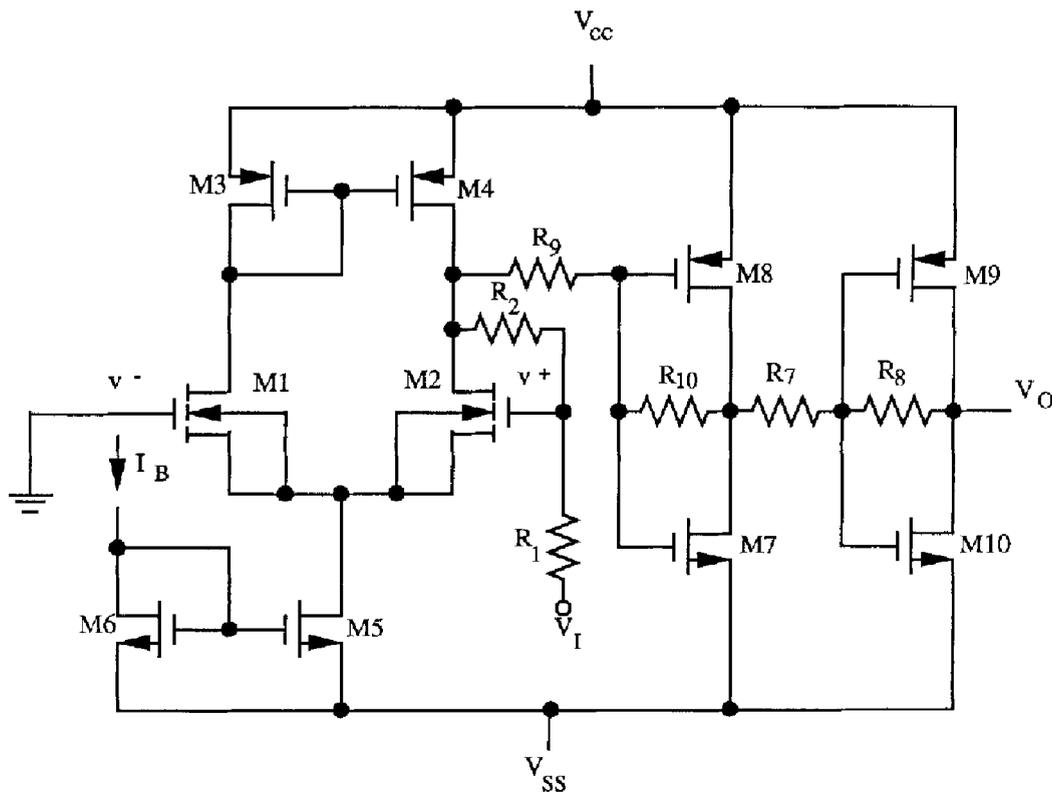


Fig. 2. A detailed CMOS circuit diagram of the sigmoid neuron.

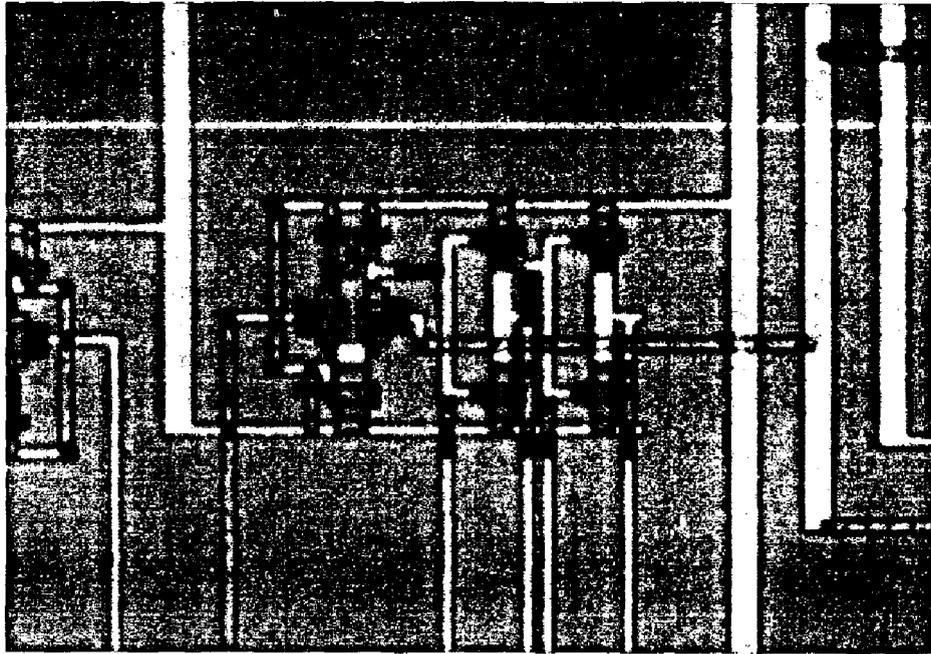


Fig. 3. The layout of the fabricated sigmoid function chip.

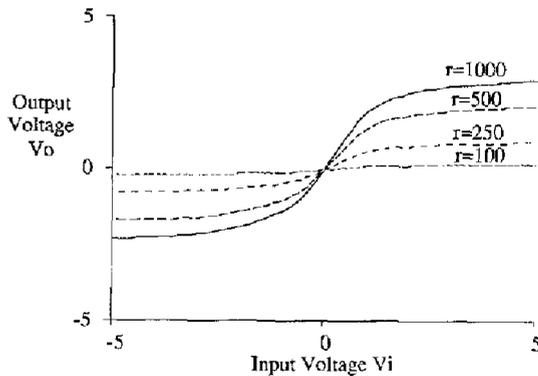


Fig. 4. The experiment results of I/O response on test chip.

the convergence time to the global minimum. The input/output function of the hysteresis McCulloch-Pitts binary neuron is given by [7]:

$$\begin{aligned}
 V_i = f(U_i) &= 1 \quad \text{if } U_i \geq \text{UTP (upper trip point)} \\
 &= 0 \quad \text{if } U_i \leq \text{LTP (lower trip point)} \\
 &\quad \text{unchanged if } \text{LTP} < U_i < \text{UTP}
 \end{aligned}
 \quad (2)$$

Figure 5 shows the digital circuit diagram of the hysteresis McCulloch-Pitts binary neuron [8]. We designed a synchronous sequential circuit using flip-flop programming to realize the hysteresis property. A 6-bit fixed point 2's complement arithmetic is used

for calculation of the input U . The 5-bit input $U(U_0, U_1, \dots, U_4)$ is compared at the 5-bit comparator with the specific value which will be equal to either LTP or UTP. A multiplexer realized by three NOR gates transmits the value of LTP or the value of UTP according to the previous state of the neuron itself and is stored in a leftmost flip-flop. The state of every neuron is updated synchronously with a clock signal labeled CK2. The 3-bit value of UTP is stored in the 3-bit shift register through the serial input line named HYS. For this purpose, another clock signal labeled CK1 is prepared.

Figure 6 shows the layout diagram of the hysteresis McCulloch-Pitts binary neuron using the MOSIS scalable CMOS/bulk VLSI technology with 2- μm rule of P-well, double-level metal. A neuron needs a $299\lambda \times 368\lambda$ layout area.

Figure 7 shows the simple asynchronous digital circuit for the hysteresis McCulloch-Pitts binary neuron. A 5-bit fixed point 2's complement arithmetic is used for the input U . For simplicity, LTP is set at 0 and UTP is set a 2. The 5-bit input $U(U_0, U_1, \dots, U_4)$ is provided with the specific value of LTP and UTP. The output V is 1 if input U is greater than 2, V is 0 if U is less than 0, and V remains unchanged otherwise. Figure 8 shows the timing diagram of the simulation result. Figure 9 shows the layout diagram of an asynchronous hysteresis McCulloch-Pitts binary neuron using the

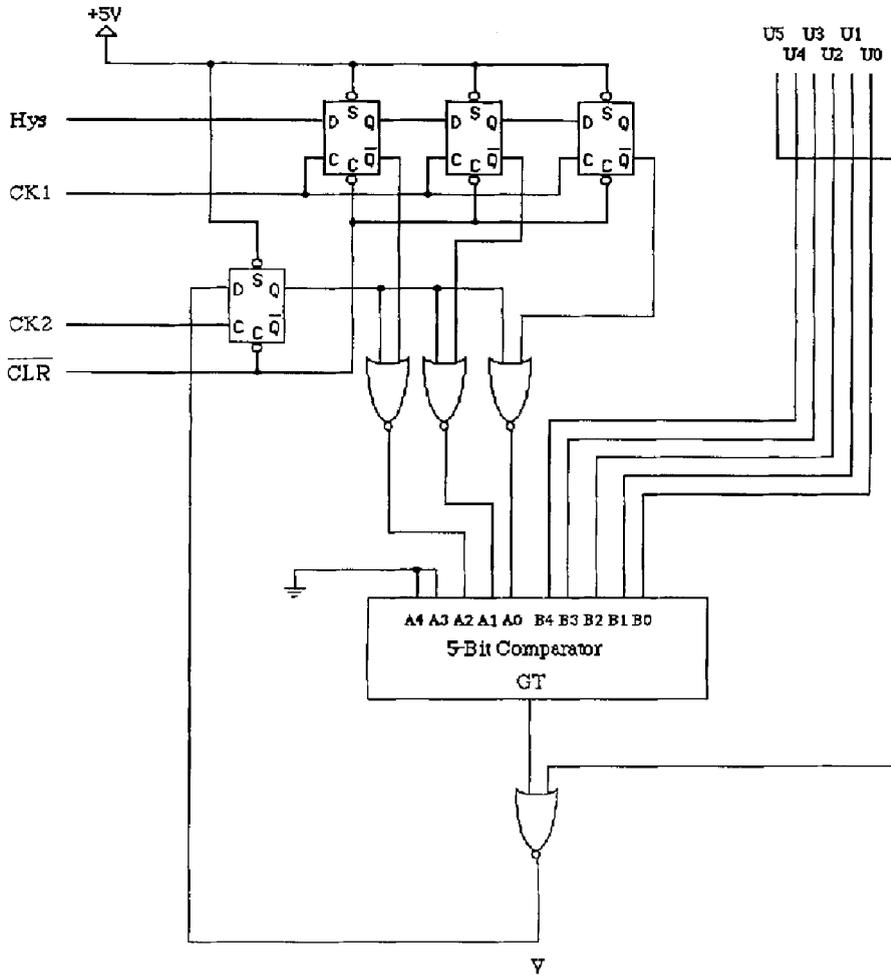


Fig. 5. Digital circuit diagram of the hysteresis McCulloch-Pitts binary neuron.

MOSIS scalable CMOS/bulk VLSI technology with 2- μm rule of P-well, double-level metal. A neuron needs a $92\lambda \times 109\lambda$ layout area.

The asynchronous implementation is close to the biological world. The advantage of the asynchronous circuit is its smaller area, and thus larger networks can be implemented in a chip. For the same complexity of the neuron architecture the asynchronous circuit requires just 10% of asynchronous silicon area.

3. Ramsey Number Problem

3.1. Problem Description [10]

The Ramsey number, $R(r, b)$ is given by the smallest value of N ; $R(r, b) = N$ where all edges of a complete graph of $N - 1$ vertices are colored either red or blue,

and red and blue subgraphs should not form any complete subgraphs of r vertices and that of b vertices respectively. A superimposed red and blue subgraph is called a Ramsey graph.

According to the latest results for the Ramsey numbers [11, 12], only seven Ramsey numbers, $R(3,3)$, $R(3,4)$, $R(3,5)$, $R(3,6)$, $R(3,7)$, $R(3,9)$, and $R(4,4)$ have been found and fixed, and the others are still unknown.

3.2. System Representation and Parallel Algorithm

We first define that V_{ijk} and U_{ijk} are the output and input of the ijk th processing element, respectively and the ijk th processing element represents the state of an edge between the i th vertex and the j th vertex with k th color. $V_{ij1} = 1$ means the ij th edge is red and $V_{ij2} = 1$ is blue.

and $V_{ijk} = 0$ respectively. The 13th edge, the 14th edge, the 23rd edge, 25th edge, and the 45th edge are red ($k = 1$), and the 12th edge, 15th edge, the 24th edge, the 34th edge, and the 35th edge are blue ($k = 2$).

The state of the system in figure 10b represents the subgraphs in figure 10c. The superimposed two-subgraphs becomes the Ramsey graph of the 5-vertex complete graph.

The hysteresis McCulloch-Pitts neuron model is used to give the value of V_{ijk} . The change of the U_{ijk} is given by the motion equation

$$\frac{dU_{ijk}}{dt} = - \frac{\partial E(V_{121}, V_{122}, \dots, V_{n-1, n1}, V_{n-1, n2})}{\partial V_{ijk}} \quad (3)$$

The energy function of the state, E is basically composed of two parts: color constraint and edge violation.

The color constraint term for the ij th edge is given by the equation

$$L[i, j, k] = \frac{1}{2} \left(\sum_{k=1}^2 V_{ijk} - 1 \right)^2 \quad (4)$$

This term is to force one and only one neuron to fire in the ij th edge. In other words, the ij th edge must be red or blue. If both two or no neuron is fired in the ij th edge, then it becomes 1, 0 otherwise.

The edge violation conditions for the ijk th edge are shown in figure 11. If the assignment of the ijk th edge

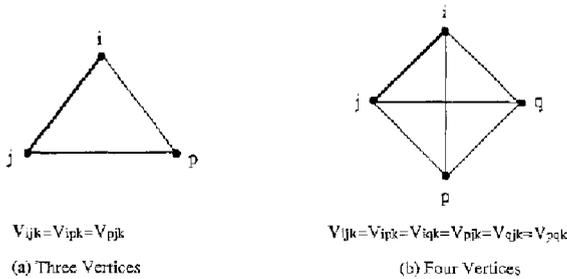


Fig. 11. Violation conditions for the ijk th edge.

to the red (blue) subgraph creates complete subgraph of $r(b)$ vertices, the ijk th edge must not be embedded in it. These violation terms $R[i, j, 1]$ for red edges and $B[i, j, 2]$ for blue are given by

$$R[i, j, 1] = \sum_{\substack{p=1 \\ p \neq i, j}}^n V_{ip1} V_{pj1} V_{ij1} \quad \text{for } r = 3 \quad (5)$$

$$R[i, j, 1] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n V_{ip1} V_{iq1} V_{pj1} V_{qj1} V_{pq1} V_{ij1} \quad \text{for } r = 4 \quad (6)$$

$$B[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n V_{ip2} V_{pj2} V_{ij2} \quad \text{for } b = 3 \quad (7)$$

$$B[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n V_{ip2} V_{iq2} V_{pj2} V_{qj2} V_{pq2} V_{ij2} \quad \text{for } b = 4 \quad (8)$$

$$B[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n \sum_{\substack{r=q+1 \\ r \neq i, j}}^n V_{ip2} V_{iq2} V_{ir2} V_{pj2} V_{qj2} V_{rj2} V_{pr2} V_{qr2} V_{ij2} \quad \text{for } b = 5 \quad (9)$$

$$B[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n \sum_{\substack{r=q+1 \\ r \neq i, j}}^n \sum_{\substack{s=r+1 \\ s \neq i, j}}^n V_{ip2} V_{iq2} V_{ir2} V_{is2} V_{pj2} V_{qj2} V_{rj2} V_{sj2} \times V_{pq2} V_{pr2} V_{ps2} V_{qr2} V_{qs2} V_{rs2} V_{ij2} \quad \text{for } b = 6 \quad (10)$$

These conditions are nonzero if the assignment of the ijk th edge to the red (blue) subgraph creates a complete subgraph of $r(b)$ vertices.

The energy functions of the ijk th processing element for finding a Ramsey graph of the n -vertex complete graph $R(r, b)$ are given by

$$E_{ij1} = \alpha \cdot L[i, j, 1] + \beta \cdot R[i, j, 1] - \gamma \cdot B[i, j, 1] \quad (11)$$

$$E_{ij2} = \alpha \cdot L[i, j, 2] - \beta \cdot R[i, j, 2] + \gamma \cdot B[i, j, 2] \quad (12)$$

where α , β , and γ are constant coefficients.

These motion equations of the ijk th processing element are given by equation (3) as follows:

$$\frac{dU_{ij1}}{dt} = -\alpha \cdot L'[i, j, 1] - \beta \cdot R'[i, j, 1] + \gamma \cdot B'[i, j, 1] \quad (11')$$

$$\frac{dU_{ij^2}}{dt} = -\alpha \cdot L'[i, j, 2] - \beta \cdot R'[i, j, 2] + \gamma \cdot B'[i, j, 2] \quad (12')$$

where

$$L'[i, j, k] = \sum_{k=1}^2 \frac{1}{V_{ij} - 1} \quad (4')$$

$$R'[i, j, 1] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \frac{1}{V_{ip1} V_{pj1}} \quad \text{for } r = 3 \quad (5')$$

$$R'[i, j, 1] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n \frac{1}{V_{ip1} V_{iq1} V_{pj1} V_{qj1} V_{pq1}} \quad \text{for } r = 4 \quad (6')$$

$$B'[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \frac{1}{V_{ip2} V_{pj2}} \quad \text{for } b = 3 \quad (7')$$

$$B'[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n \frac{1}{V_{ip2} V_{iq2} V_{pj2} V_{qj2} V_{pq2}} \quad \text{for } b = 4 \quad (8')$$

$$B'[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n \sum_{\substack{r=q+1 \\ r \neq i, j}}^n \frac{1}{V_{ip2} V_{iq2} V_{ir2} V_{pj2} V_{qj2} V_{rj2} V_{pq2} V_{pr2} V_{qr2}} \quad \text{for } b = 5 \quad (9')$$

$$B'[i, j, 2] = \sum_{\substack{p=1 \\ p \neq i, j}}^n \sum_{\substack{q=p+1 \\ q \neq i, j}}^n \sum_{\substack{r=q+1 \\ r \neq i, j}}^n \sum_{\substack{s=r+1 \\ s \neq i, j}}^n \frac{1}{V_{ip2} V_{iq2} V_{ir2} V_{is2} V_{pj2} V_{qj2} V_{rj2} V_{sj2} \times V_{pq2} V_{pr2} V_{ps2} V_{qr2} V_{qs2} V_{rs2}} \quad \text{for } b = 6 \quad (10')$$

The first term on the right side of eq. (11') encourages the sum of the output of the ijk th processing element to be one so that the ijk th edge is colored red or blue. The second term encourages the output of the ij th processing element to be zero if the ij th edge creates a complete subgraph of r vertices. The third term encourages the output of the ij th processing element to be nonzero if the ij th edge creates a complete subgraph of b vertices.

In addition, we add the following hill-climbing term to each motion equation:

$$\omega \cdot H[i, j, k] = \omega \cdot h \left(\sum_{k=1}^2 V_{ijk} \right) \quad (13)$$

where $h(x) = 1$ if $x = 0$, 0 otherwise and ω is a constant coefficient. This term also plays an important role to increase the frequency of the state of the system to converge to the global minimum [3].

In this paper, the synchronous parallel system was simulated on a Macintosh Powerbook170, although the algorithm is executable in a parallel machine. The synchronous parallel system can be performed on maximally $n(n - 1)$ processors for finding a Ramsey graph of the n -vertex complete graph. The state of $n(n - 1)$ processing elements can be updated by using equations (11'), (12'), and (13) with the first-order Euler method.

3.3. Results and Discussion

The following five problems were solved by the algorithm.

1. Problem 1: finding the Ramsey graph of 5-vertex complete graph for $R(3,3)$
2. Problem 2: finding the Ramsey graph of 8-vertex complete graph for $R(3,4)$
3. Problem 3: finding the Ramsey graph of 13-vertex complete graph for $R(3,5)$
4. Problem 4: finding the Ramsey graph of 17-vertex complete graph for $R(3,6)$
5. Problem 2: finding the Ramsey graph of 17-vertex complete graph for $R(4,4)$

Figures 12, 13, 14, and 15 show the simulation results for problems 2, 3, 4, and 5 respectively. The algorithm found several Ramsey graphs from the different initial values of U_{ijk} . Table 1 shows the average number of iteration steps and the frequency of the convergence to solutions within 700 iteration steps, where 100 simulation runs were performed for each one of five problems. In every simulation run, the different

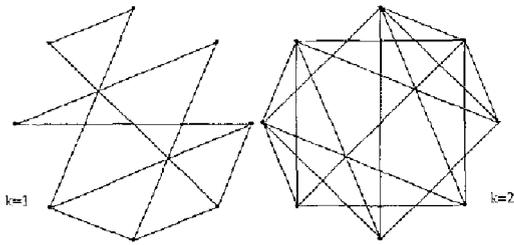


Fig. 12. A Ramsey graph with eight vertices for $R(3,4)$.

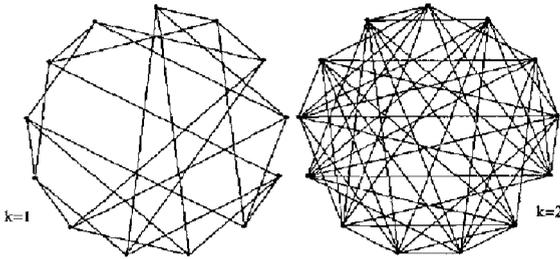


Fig. 13. A Ramsey graph with 13 vertices for $R(3,5)$.

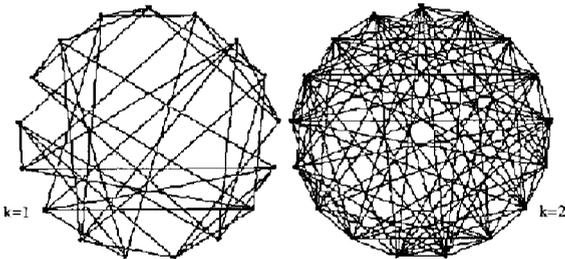


Fig. 14. A Ramsey graph with 17 vertices for $R(3,6)$.

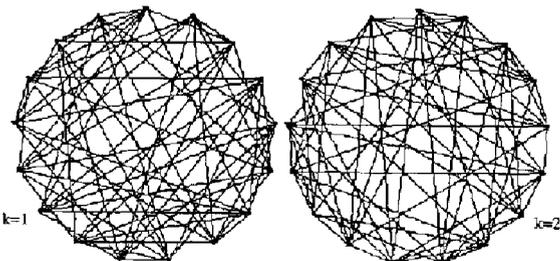


Fig. 15. A Ramsey graph with 17 vertices for $R(4,4)$.

initial values were given. Figure 16 shows the relationship between the frequency and the number of iteration steps to solutions in problems 1 and 3. The simulation results showed the consistency of the algorithm. The primary goal of finding the Ramsey graphs using our parallel algorithm was successfully achieved in terms of the computation time and the solution quality.

Table 1. Summary of simulation results.

Problem no.	Average number of iteration steps	Convergence frequency to solutions
1. $R(3,3)$	28.7	100%
2. $R(3,4)$	101.6	100%
3. $R(3,5)$	248.9	85%
4. $R(3,6)$	373.4	22%
5. $R(4,4)$	361.0	7%

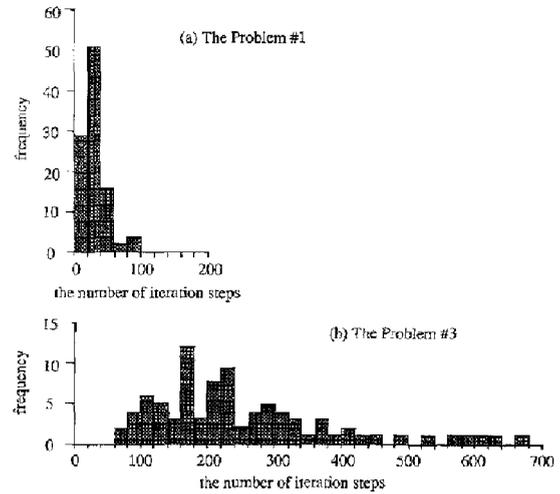


Fig. 16. Relationship between the frequency and the number of iteration steps.

4. Conclusion

Digital/analog circuits of the hysteresis McCulloch-Pitts neuron are presented in this paper. It can be used to solve a variety of NP-complete problems. Based on our estimation several thousands of asynchronous neurons can be embedded in a single chip based on the CMOS technology. The asynchronous implementation is close to the biological world. The advantage of the asynchronous circuit is its smaller area, and thus larger networks can be implemented in a chip. For the same complexity of the neuron architecture the asynchronous circuit requires just 10% of synchronous silicon area. The advantage of the fully digital implementation is good expandability to larger network. In addition, single circuits can be joined together to form very large networks. The advantages of an analog circuit are simpler basic blocks and smaller area. However, the noise immunity is worse than digital circuits.

The proposed parallel algorithm also finds Ramsey graphs of several complete graphs with different Ramsey numbers. The algorithm uses $n(n - 1)$ processing

elements for finding a Ramsey graph of an n -vertex complete graph. The simulation result of our algorithm with silicon neural network shows that our approach is very promising for finding unknown Ramsey numbers.

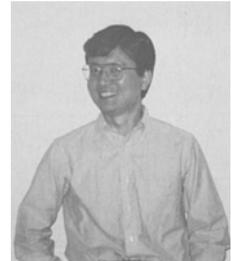
References

1. C. Mead, *Analog VLSI and Neural Systems*, Addison-Wesley: Reading, MA, 1989.
2. Y. Takefuji, "Analog components for neural networks," Course Note for Dept. of EEAP at Case Western Reserve University, 1989-1991.
3. Y. Takefuji and K.C. Lee, "Artificial neural networks for four-coloring map problem and K-colorability problems," *IEEE Trans. Circuits Syst.*, Vol. 38, No. 3, pp. 326-333, 1991.
4. J.J. Hopfield and D.W. Tank, "Neural computation of decisions in optimization problems," *Biological Cybernetics*, Vol. 52, pp. 141-152, 1985.
5. D.W. Tank and J.J. Hopfield, "Simple neural optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circuit Syst.*, Vol. CAS-33, No. 5, pp. 533-541, 1986.
6. Y. Takefuji and K.C. Lee, "A near-optimum parallel planarization algorithm," *Science*, Vol. 245, pp. 1221-1223, 1989.
7. P.E. Allen and D.R. Holberg, *CMOS Analog Circuit Design*, Holt, Rinehart, Winston: New York, 1987.
8. Y. Takefuji and K.C. Lee, "An artificial hysteresis binary neuron: a model suppressing the oscillatory behaviors of neural dynamics," *Biological Cybernetics*, Vol. 64, pp. 353-356, 1991.
9. T. Kurokawa, K.C. Lee, Y.B. Cho, Y. Takefuji, and N. Fonabiki, "CMOS layout design of the hysteresis McCulloch-Pitts neuron," *Electron. Lett.*, Vol. 26, p. 25, 1990.
10. G. Polya, R.E. Tarjan, and D.R. Woods, *Notes on Introductory Combinatorics*, Birkhauser: Boston, 1983.
11. R.L. Graham, B.L. Rothschild, and J.H. Spencer, *Ramsey Theory*, Wiley: New York, 1980, 1990.
12. S.P. Radziszowski and D.L. Kreher, "Search algorithm for Ramsey graphs by union of group orbits," *J. Graph Theory*, Vol. 12, No. 1, pp. 59-72, 1988.



Yong Beom Cho received the B.S. degree in electronics engineering from Kyungpook National University, Taegu, Korea, in 1981, the M.S. degree in electrical engineering and computer engineering from University of South Carolina, Columbia, SC, in 1988, and the Ph.D. degree in electrical engineering and applied physics from Case Western Reserve University, Cleveland, OH, in 1992. He has been with J.W. Harley, Inc., Twinsburg, OH, since 1990, working in areas of A.I. applications for power engineering. From 1988 to 1990, he

was a research assistant with the Department of Electrical Engineering and Applied Physics, Case Western Reserve University, working in areas of artificial neural network applications for optimization problems. His research interests include neural networks, expert system, and VLSI design.



Kazuhiro Tsuchiya received his B.S. (1983) in chemical engineering and M.S. (1985) in material chemistry from Tohoku University (Japan). His research interests focus on neural network parallel computing for solving real-world problems and VLSI applications and silicon architecture. He has worked for Fuji Electric Co., Ltd. in Japan since 1985.



Yoshiyasu Takefuji has been an associate professor on the faculty of environmental information at Keio University since 1992 and has been on the electrical engineering faculty at Case Western Reserve University since 1988. Before joining Case, he taught at the University of South Florida and the University of South Carolina. He received his B.S. (1978), M.S. (1980), and Ph.D. (1983) in electrical engineering from Keio University (Japan). His research interests focus on neural network parallel computing for solving real-world problems. He is interested in VLSI applications and silicon architecture. He received the National Science Foundation/Research Initiation Award in 1989 and is an NSF advisory panelist. A member of the IEEE Computer Society, ACM, International Neural Network Society, and American Association for the Advancement of Science, he received the Information Processing Society of Japan's best paper award in 1980. He has written a book entitled *Neural Network Parallel Computing*, published by Kluwer, and has co-authored two books, *Digital Circuits* (Ohm-Sha Publishers) in 1984 and *Neural Network Computing* (Baifukan Publishers) in 1992. He was an editor of the *Journal of Neural Network Computing* and is an associate editor of *IEEE Transactions on Neural Networks and Neurocomputing*. He is the guest editor of the *Journal Analog Integrated Circuits and Signal Processing* in the special issue on analog VLSI neural networks, and the guest editor of *Neurocomputing* in the special issue on neural-network-based optimization. He has published more than 100 papers.