D1-MACA 기반의 두 클래스 패턴 분류기

황윤희* · 최언숙** · 조성진*＋

D1-MACA based Two-Class Pattern Classifier

Yoon-hee Hwang* · Un-sook Choi** · Sung-jin Cho*

요 약

이 논문에서는 주어진 패턴 집합을 두 개의 분할된 클래스로 분류하는 분류기로써 D1-MACA (Depth 1 Multiple Attractor Cellular Automata)를 제안한다. 이 때 메모리량을 최소화 할 수 있는 방법으로 attractor의 수가 2개 되게 D1-MACA를 구성할 수 있는 패턴 집합의 조건을 분석하고, 분류기로써의 D1-MACA를 구성하는 방법을 부분공간의 개념을 이용하여 효율적으로 구성한다.

ABSTRACT

Two with the shunt which classifies with the class which is divided D1-MCA(Depth 1 Multiple Attractor Cellular Automata) proposes pattern gathering which comes to give from this dissertation. This time attractor valence 2 makes become with the method will be able to minimize the memory quantity the condition of pattern gathering will be able to compose D1-MACA analyzes classification crossroad bitterly the method which composes D1-MACA the concept of subspace and uses composes efficiently.

키워드

MACA, Attractor, Patteren Classifier, Partial, D1-MACA

Ⅰ. 서 론

셀룰라 오토마타(Cellular Automata, 이하 CA)는 간단하고 규칙적이며 작은 단위로 확장 연결이 가능한 구조이기 때문에 하드웨어 구현에 적합하여 LFSR의 대안으로 제한되었다. 이러한 CA는 상태 전이 그래프에서 모든 상태가 사이클에 놓이는 그룹 CA와 그렇지 않은 비그룹 CA로 나뉜다. 그룹 CA는 테스트 패턴 생성, 의사 난수열 생성, 오류정정부호의 설계 등에 응용되면서 활발하게 연구되어 왔다[1][2]. 또한 비그룹 CA는 이미지 압축, 해쉬 함수, 부울 방정식의 해법 등에 응용되면서 최근 연구가 활발히 이루어지고 있다[3][4]. 특히 비그룹 CA의 상태 전이 그래프는 길이가 1인 사이클을 루트(root)로 갖는 분할된 집합으로 구성된다. 이것은 자연스러운 분류기(classifier)의 형태를 가진다. 분류의 문제는 데이터베이스 시스템에서 기록의 그룹화, VLSI 회로에서의 결합 찾는 것, 이미지 프로세싱과 같은 컴퓨터 과학에 중요한 역할을 한다. [5]에서는 주어진 패턴 집합이 두 개의 클래스로 분류되는 패턴들을 분류할 수 있는 분류기를

* 부경대학교
** 동명대학교
+ : 교신저자

접수일자 : 2008. 09. 22
심사완료일자 : 2008. 11. 13

구성하는 방법으로 D1-MACA (Depth 1 Multiple Attractor CA)를 사용하고 있다. D1-MACA에 기반한 분류기는 기존의 방법보다 저장해야 하는 메모리량이 적은 장점을 가지고 있으나 [5]에서는 이러한 CA를 구성하는 방법이 다소 복잡하고 어떤 패턴의 경우에 분류기로써의 역할을 못하고 있다.

이 논문에서는 주어진 패턴 집합이 두 개의 클래스로 분류되는 패턴들을 분류할 수 있는 분류기를 여러 개의 attractor를 갖고 depth가 1인D1-MACA(Depth 1 Multiple Attractor CA)를 부분공간의 개념을 이용하여 효율적으로 구성한다. 또한 메모리량을 최소화 할 수 있는 방법으로 attractor의 수가 2개 되게 D1-MACA를 구성할 수 있는 패턴 집합의 조건을 분석한다.

II. 셀룰라 오토마타의 기본성질

이 절에서는 본 논문에서 사용되는 CA의 용어와 기본성질에 대해 간단히 언급한다.

attractor : 비그룹 CA의 상태전이 그래프에서 순환 상태들 중 사이클의 길이가 1인 상태를 말한다.

depth : 비그룹 CA의 상태전이 그래프에서 임의의 도달불가능한 상태에서 가장 가까운 순환상태로 가는데 걸리는 최소의 단계 수를 말한다.

<정리 1 [6]> 상태 전이 행렬 T를 갖는 n-셀 MACA에 대하여 $T\oplus I$의 차원이 r일 때 attractor의 수는 2^{n-r}개이다.

<정리 2 [6]> 2^m개의 attractor를 갖는 CA에서 각 attractor에서 2^m개의 의사 전수 패턴(pseudo-exhaustive pattern)을 갖는 비트 위치가 존재한다.

<예제 3> 4-셀 CA의 상태 전이 행렬이 다음과 같을 때 $T\oplus I$의 차원은 2차이다. 따라서 2^2개의 attractor를 가짐을 알 수 있다.

$$T=\begin{pmatrix}1&0&0&0\\0&1&1&0\\0&1&1&0\\0&0&0&1\end{pmatrix},\quad T\oplus I=\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&0\end{pmatrix}$$

이 CA의 상태 전이 그래프를 그리면 그림 1과 같음을 확인할 수 있다.

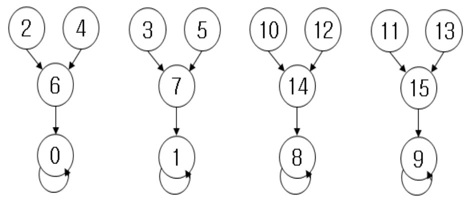

그림 1. MACA의 상태 전이 그래프
Fig. 1 MACA의 상태 전이 그래프

여기서 attractor 집합은 $\{0,1,8,9\}$로 2-비트의 의사 전수 패턴을 첫 번째 비트와 네 번째 비트를 택함으로써 다음과 같이 생성할 수 있다.

0000
0001
1000
1001

depth가 1인 MACA를 D1-MACA라 하는데 이러한 CA는 분류기를 디자인 하는데 유용하다. 왜냐하면 D1-MACA는 모든 도달 불가능한 상태들이 CA를 돌렸을 때 한 번만에 attractor에 도달할 수 있기 때문에 분류 시 필요한 시간을 줄일 수 있다.

<정리 4 [5]> depth가 d인 MACA에 대하여 D1-MACA가 존재한다.

depth가 d인 MACA의 상태 전이 행렬이 T일 때 D1-MACA의 상태 전이 행렬은 T^d가 된다. 예를 들어 예제 3에서 depth가 2이므로 T^2을 상태 전이 행렬로 둔다면 그림 2의 상태 전이 그래프를 갖는 D1-MACA가 된다.

$$T^2 = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix}$$

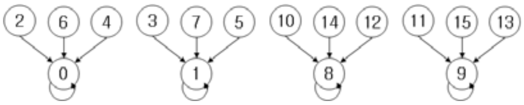

그림 2. D1-MACA의 상태 전이 그래프
Fig. 2 Pattern shunt of D1-MACA bases

III. D1-MACA 기반의 패턴 분류기

m개의 attractor를 갖는 n-비트 MACA는 분류기로 볼 수 있다. 각 attractor에 대한 트리의 상태들은 m개의 분할된 분류로 나눠진다. 그림 3에서와 같이 의사 전수 패턴을 생성하는 비트 위치의 값들 만으로 그 패턴의 클래스를 구별할 수 있다.

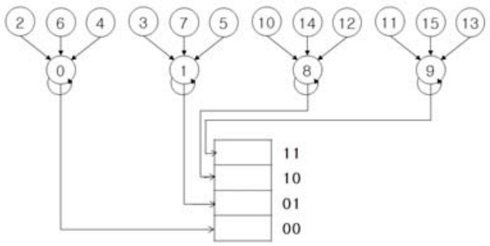

그림 3. MACA 기반의 분류기
Fig. 3 Shunt of MACA bases

주어진 패턴 집합 P가 단지 두 개의 분할된 패턴 분류 P_1와 P_2로 구성된다고 하자. 즉, $P = P_1 \cup P_2$이고 $P_1 \cap P_2 = \varnothing$이 성립한다. 만약 분류기가 P를 분류한다면, 주어진 두 패턴 $x \in P_1, y \in P_2$은 다른 트리에 놓이게 된다. 따라서 만약 T가 이러한 분류를 수행하는 D1-MACA의 상태 전이 행렬이면 다음의 두 관계를 만족하여야 한다.

관계 1>P_1의 임의의 원소 x와 P_2의 임의의 원소 y에 대하여 $T \bullet (x) \neq T \bullet (y)$이 성립하여야 한다. 즉, $T \bullet (x \oplus y) \neq O$이다.

관계 2> depth가 1이므로 $T^2 = T$이여야 한다. 즉, $T(T \oplus I) = O$이 성립하여야 한다.

위 두 관계를 만족하는 D1-MACA를 구성하는 것에는 각 패턴이 어떻게 구성되는 상황에 의하여 2개 이상의 attractor를 가지게 된다. 만약 P_1의 원소가 D1-MACA의 a_1개의 트리에 놓여있고, P_2의 원소가 D1-MACA의 a_2개의 트리에 놓여있다고 하자. 그 때 의사 전수 패턴을 생성하는 m비트가 있다고 가정한다면 메모리 량 다음과 같다[5].

$$M_{CA} = \min(a_1, a_2) \times m + 1$$

이 때 메모리를 가장 적게 차지하는 분류기를 가장 선호하게 되는데 이는 attractor가 두 개로만 구성된 D1-MACA일 경우가 된다.

주어진 패턴 집합에 대하여 attractor가 두 개로만 구성된 D1-MACA를 구성할 수 있는 조건은 다음과 같다.

조건 1> 패턴 집합 P의 두 개의 분할된 P_1, P_2들 중에서 적어도 하나는 0을 가지지 않고, 모든 원소들이 일차독립인 집합이어야 한다.

조건 2> 조건 1을 만족하는 집합을 제외한 남은 나머지 집합의 원소들의 일차결합들이 조건 1을 만족하는 집합의 원소이면 안된다.

조건 3> 조건 1을 만족하는 집합의 원소들 중에서 임의의 홀수 개의 일차결합들은 조건1을 만족하는 집합을 제외한 남은 나머지 집합의 원소이면 안된다.

조건 4> 조건 1을 만족하는 집합의 원소들 중에서 임의의 짝수 개의 일차결합들은 그 집합의 원소가 되어서는 안된다.

위의 4가지 조건을 만족하지 않는다면 attractor의 수는 2를 넘게 된다.

<**예제** 5> 주어진 4-비트 패턴 집합 P는 다음 과

같은 분할된 두 패턴 분류 P_1, P_2 로 구성되어 있다고 가정하자.

$$P = \{0,1,2,5,6,7,10,11,15\}$$
$$P_1 = \{0,1,6,7,15\}$$
$$P_2 = \{2,5,10,11\}$$

이러한 패턴 집합은 P_2의 원소들이 0을 갖지 않고 일차독립이므로 조건 1을 만족한다. P_1의 원소들의 일차결합들이 P_2의 원소가 아니므로 조건 2를 만족한다. 유사한 방법으로 위의 4가지 조건을 모두 만족한다 따라서 이 패턴을 분류하기 위한 분류기로 attractor가 두 개인 D1-MACA를 구성할 수 있다. 상태 전이 행렬 T를 구성하기 위하여 다음을 구하여 보자.

관계 1을 만족하기 위하여 $P_1 \oplus P_2 = \{2,3,4,5,10,11,12,13\}$의 모든 원소는 절대 attractor가 0인 트리의 원소여서는 안 된다.

step 1) $P_1 \oplus P_2$의 원소 중 일차 독립인 벡터의 수는 4개이다.

step 2) $(P_1 \oplus P_2)^c = \{0,1,6,7,8,9,14,15\}$의 일차독립인 원소를 이진화 했을 때 이 중에서 각 비트에 1이 개수가 가장 적은 것을 찾으면 다음과 같다.

0	0000		
1	0001		
6	0110		0001
7	0111	⇨	0110
8	1000		1000
9	1001		
14	1110		
15	1111		

stpe 3) step 2에서 구한 원소들은 모두 attractor가 0인 트리로 위치하게 D1-MACA를 구성하면 다음과 같다.

$$T = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&1&1&0 \\ 0&0&0&0 \end{pmatrix}$$

그림 4는 주어진 패턴을 분류하는 T를 상태 전이 행렬로 갖는 D1-MACA의 상태 전이 그래프이다.

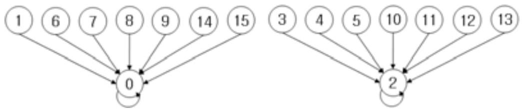

그림 4. P_1과 P_2를 분류한 D1-MACA
Fig. 4 D1-MACA where classifies P_1 and P_2

여기서 T는 다음과 같이 구성하여도 된다.

$$T = \begin{pmatrix} 0&0&0&0 \\ 0&1&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}, \quad T = \begin{pmatrix} 0&0&0&0 \\ 0&1&1&0 \\ 0&1&1&0 \\ 0&0&0&0 \end{pmatrix}$$

만약 4가지 조건 중 어느 하나도 맞지 않는다면 attractor가 2개 이상인 D1-MACA를 고려하여야 한다.

<예제 6> 주어진 4-비트 패턴 집합 P는 다음 과 같은 분할된 두 패턴 클래스 P_1, P_2 로 구성되어 있다고 가정하자.

$$P = \{0,1,2,3,7,8,9,11\}$$
$$P_1 = \{7,9,11\}$$
$$P_2 = \{0,1,2,3,8\}$$

이 패턴 집합은 위의 4가지 조건을 모두 만족시키지 못한다. 따라서 attractor 수가 4개 이상이여야 한다.

step 1) $P_1 \oplus P_2 = \{1,3,4,5,6,7,8,9,10,11,15\}$의 원소 중 일차 독립인 벡터의 수는 4개이다.

step 2) $(P_1 \oplus P_2)^c = \{0,2,12,13,14\}$는 부분공간이 아니다. 따라서 부분공간을 $\{0,2,12,14\}$로 구성한다. 이 공간의 차원은 2이다. 따라서 P_1은 attractor가 0가

아닌 트리에 배열되어야 하고 밑줄 친 비트에서 의사 전수 패턴이 생성되므로 P_1은 모두 같은 트리에 존재함을 알 수 있다.

$$
\begin{array}{c}
0\underline{11}1 \\
1\underline{00}1 \\
1\underline{01}1 \\
0\underline{10}1
\end{array}
$$

step 3) $(P_1 \oplus P_2)^c$의 원소 중 attractor가 0인 트리에 존재하지 않는 원소 13을 이용하여 하나의 트리를 구성하면 MACA의 성질[7]에 의하여 13을 포함하는 트리는 $15(=13\oplus 2), 1(13\oplus 12)$, $3(=13\oplus 14)$을 원소로 갖게 된다.

step 4) 남아있는 P_2 원소 중 8을 이용하여 step 3과 유사한 방법으로 나머지 트리를 구성한다.

따라서 다음과 같이 분류될 수 있는 분류기를 D1-MACA로 구성하면 된다.

표 1. 패턴 분류
Table. 1 Pattern classification

P_2			P_1
$0\underline{00}0$	0001	1000	1001
0010	$0\underline{01}1$	1010	1011
1100	1101	$0\underline{10}0$	0101
1110	1111	0110	$0\underline{11}1$

step 5) attractor는 $(T\oplus I)$의 영공간이므로 부분공간이어야 하므로 의사 전수 패턴이 나오는 셀이 존재하여야 한다. 표 1에서 밑줄 친 부분에서 2-비트 의사 전수 패턴이 나오므로 0, 3 4 7을 attractor로 두게 되면 상태 전이 행렬 T은 다음과 같다.

$$T=\begin{pmatrix} 0\,0\,0\,0 \\ 1\,1\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,0\,1 \end{pmatrix}$$

IV. 결 론

이 논문에서는 주어진 패턴 집합이 두 개의 클래스로 분류되는 패턴들을 분류할 수 있는 분류기를 여러 개의 attractor를 갖고 depth가 1인 D1-MACA(Depth 1 Multiple Attractor CA)를 부분공간의 개념을 이용하여 효율적으로 구성하였다. 또한 메모리량을 최소화할 수 있는 방법으로 attractor의 수가 2개 되게 D1-MACA를 구성할 수 있는 패턴 집합의 조건을 분석하였다.

참고 문헌

[1] P.H. Bardell, “Analysis of cellular auto- mata used as pseudorandom pattern generators”, Proc. IEEE int. Test. Conf., pp.762～767, 1990.

[2] S.U. Guan and S.K. Tan, “Pseudorandom number generation with self-programmable cellular automata”, IEEE Transaction on computer-aided design of integrated circuits and systems, Vol. 23, No. 7, pp.1095-1101, 2004.

[3] S. Chakraborty, D.R. Chowdhury and P.P. Chaudhuri, “Theory and application of non-group cellular automata for synthesis of easily testable finite state machines”, IEEE, Trans, Computers, Vol. 45, pp.769-781, 1996.

[4] S.J. Cho, U.S. Choi, H.D. Kim, Y.H. Hwang and J.K. Kim, “Analysis of 90/150 two predecessor nongroup cellular automata”, LNCS 5191, pp.128-135, 2008.

[5] S. Chattopadhyay, S. Adhikari, S. Sengupta and M. Pal, “Highly regular, modular, and cascadable design of cellular automata- based pattern classifier”, IEEE Transactions on VLSI systems, Vol. 8, No. 6, pp.724-734, 2000.

[6] P.P. Chaudhuri, D.R. Chowdhury, S. Nandy and Chattopadhyay, “Additive cellular automata theory and application”, 1, IEEE Computer Society Press, California, 1997.

[7] S.J. Cho, H.D. Kim and U.S. Choi, “Behavior of complemeted cellular automata derived from a linear cellular automata”, Mathematical and Computer Modelling, Vol. 36, pp.979-986, 2002.

저자 소개

황윤희(Yoon-hee Hwang)

2002년 2월 : 부경대학교 통계학과 학사
2004년 2월 : 부경대학교 응용수학과 석사
2008년 8월 : 부경대학교 정보보호학과 박사
※ 관심분야 : 셀룰라 오토마타론, 정보보호, 유한체, 컴퓨터 구조론, VLSI

최언숙(Un-sook Choi)

1992년 성균관대학교 산업공학과 학사
2000년 부경대학교 응용수학과 석사
2004년 부경대학교 응용수학과 박사
2004년~2006년 2월 영산대학교 자유전공학부 단임교수
2006년 3월~현재 동명대학교 멀티미디어공학과 전임강사
※ 관심분야 : 셀룰라 오토마타론, 정보보호, 부호이론, 컴퓨터 구조론, VLSI

조성진(Sung-jin Cho)

1979년 2월 : 강원대학교 수학교육과 학사
1981년 2월 : 고려대학교 수학과 석사
1988년 2월 : 고려대학교 수학과 박사
1988년 ~ 현재 : 부경대학교 수리과학부 교수
※ 관심분야 : 셀룰라 오토마타론, 정보보호, 부호이론, 컴퓨터 구조론, VLSI