

# 룰 클러스터링에 의한 실시간 적응행동 분류자 시스템

## Classifier System for Real time Adaptive Behavior Based on Rule Clustering

황철민, 김지윤, 김현영, 심귀보

중앙대학교 전자전기공학부

Chul-Min Hwang, Ji-Yoon Kim, Hyun-Young Kim, and Kwee-Bo Sim

School of Electrical and Electronic Engineering, Chung-Ang University

E-mail : kbsim@cau.ac.kr

### ABSTRACT

기계학습의 한 종류인 분류자 시스템은 간단한 문제에 대하여 실시간 처리와 온라인 학습이 가능하다. 그러나 복잡한 환경에서는 빠른 적응이 힘들다. 본 논문에서는 복잡한 환경에서 분류자 시스템의 적응 성능을 개선함으로써 실시간이 가능하도록 전체 환경을 분류하고 각기 다른 룰 셋을 이용하는 룰 클러스터링에 의한 분류자 시스템을 제안한다. 환경을 상황에 따라 나눔으로써 전체 환경이 변화하였을 경우 각 상황에 따른 변화에 대해서만 추가적으로 학습함으로써 탐색 공간을 줄여 학습 시간을 감소시킨다. 제안한 시스템은 분류자 시스템 중 ZCS를 이용하여 로봇축구 시스템에 적용하여 기존의 방법과 그 성능을 비교 검토한다.

**Key words :** 유전자 알고리즘, 온라인 학습, 분류자 시스템

## 1. 서 론

분류자 시스템은 GBML(Genetic Based Machine Learning)의 한 종류로써 Holland에 의해서 제안되었다[1]. 분류자 시스템의 목표는 전체의 룰들 중에서 환경에 대하여 서로 협조행동을 하는데 필요한 유용한 룰들을 찾아내는 것이다. 룰들의 strength를 할당하기 위하여 'bucket brigade algorithm'을 사용하며, 할당 받은 strength에 근거하여 유용하지 않은 룰을 없애고 새롭고 유용한 룰을 탐색한다. 룰 탐색을 위하여 GA연산을 수행하여 새로운 룰을 생성하도록 함으로서 좋은 룰로의 수렴성을 보장 받는다. 룰의 구성을 보면 조건부와 행동부, strength로 구성되어 있는데 조건부와 행

동부의 길이가 같게 구성되어있어 행동부에서 다른 룰을 호출할 수 있게 되어있다[2].

Wilson이 제안한 ZCS는 기존의 분류자 시스템의 개념을 유지한 상태에서 구조를 간략화 함으로서 문제에 대한 모델링을 쉽게 하였다. 또한 룰의 조건부와 행동부의 길이에 대한 조건을 없애고, 행동부에 의해서 다른 룰들이 match되는 개념을 없앴다[3].

로봇축구와 같이 동적이면서 여러 가지 상황이 뒤섞여 있는 시스템에서의 학습의 경우 하나의 규칙 베이스로 학습을 하는 것에는 한계가 있다. 좋은 규칙들을 분류하기 위하여 더 많은 시행착오가 필요하게 되고, 결국 수렴하기 위하여 더 많은 시간이 필요하게 된다. 이 문제의 해결을 위하여 각 상황을 나누고 서로 다른 규칙집합을 가지고 학습함으로써 상황의 변화에 따라 적절히 대응할 수 있는 시스템을 구성 할 필요성이 있다.

이 논문에서는 ZCS (Zeroth Level Classifier System)을 하나의 룰 셋을 이용한

본 연구는 과학기술부의 뇌신경정보학연구사업의 '뇌정보처리 메커니즘에 기반한 인간행동 시스템연구'의 연구비 지원으로 수행되었습니다. 연구비 지원에 감사드립니다.

방법과 환경에 따라 분류된 멀티 룰 셋에 기반한 방법을 로봇 축구 시스템에 적용하여 온라인 학습함으로써 각각의 성능을 비교 평가한다.

## II. ZCS의 기본 개념

ZCS은 검지기, 룰 베이스, 효과기로 구성된다. 검지기는 환경으로부터 입력을 받아 시스템에서 사용되는 코드로 변환을 시킨다. 검지기에서 변환된 코드를 입력으로 하여 룰 베이스로 넘겨준다. 룰 베이스는 "IF.. THEN.."문과 문장의 strength 부분으로 구성된 분류자들의 집합으로 구성된다. "IF"뒤에 오는 부분은 조건부로서 검지기에서 입력되는 코드와 같은 길이로 구성되며, 검지기의 입력과 같은 코드 외에 "don't care symbol"인 "#"을 추가적으로 사용한다. "#"은 "0"이나 "1"중에서 어떤 것이라도 대신 사용될 수 있음을 나타낸다. "THEN"의 뒤에 오는 부분은 행동부로서 어떤 행동을 할 것인가에 관한 내용이 코드화 되어 구성되어 있다. 효과기는 행동부의 코드화 된 내용에 따라 환경에 출력을 내보내게 된다. 이때 환경에 따라 현재의 규칙들을 평가하고, 그에 따라 나쁜 규칙들을 제거하고 적합한 규칙들을 생성하는 작업은 룰 베이스에서 이루어진다. 룰 베이스에서는 implicit bucket brigade algorithm과 환경으로부터의 보상값을 이용하여 각각의 분류자들의 strength를 평가하고, 이 평가된 strength를 이용하여 GA연산을 수행함으로써 유용하지 않은 규칙을 제거하고, 유용한 규칙들을 탐색한다[1].

## III. 룰 클러스터링에 의한 ZCS

검지기와 효과기의 구조는 ZCS의 구조와 동일하다. 가장 큰 차이는 룰 베이스에 있는데 룰 베이스의 룰 셋이 하나로 이루어진 것이 아니라 환경이 분류된 수만큼 존재하고 각각 분류된 환경에 따라 다른 룰 셋을 사용한다. 따라서 각각 다른 룰 셋을 이용하고, 각 분류된 상황에 맞게 학습하게 된다. 그림1은 전체적인 구조를 나타낸다.

그림1에서 동작순서는 다음과 같다.

- 환경으로부터 센서 입력을 받아 '효과기'로 입력이 된다.
- 효과기에서 환경으로부터 추출한 코드를 '룰 셋 선택기'와 '룰 베이스'에 입력한다. '룰 셋 선택기'에서 어떤 '룰 셋'을 사용할지를 결정하여 '룰 셋 집합'으로 보내준다.
- 결정된 '룰 셋'에서 '효과기'에서의 입력과 비교하여 'Match Set'을 구성한

다.

- 'Match Set'에서 strength를 이용하여 'Action Set'을 결정한다.
- 과거에 사용된 'Old Action Set'에게 'bid'을 지불한다.
- 'action set'의 행동부가 '효과기'로 전달된다.
- 효과기에서 코드값에 따라 행동을 한다.
- 일정 기간 행동을 한 후 환경의 변화에 따라 'action set'에 'reward'를 제공한다.

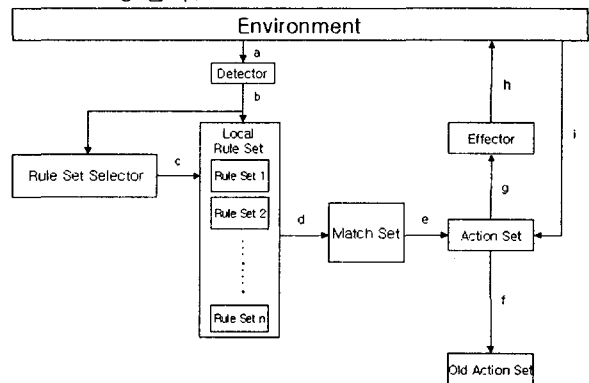


그림 1. 멀티 룰 셋에 기반한 ZCS의 구조

## IV. Simurosot을 이용한 적용

본 논문의 simulator로 사용된 Simurosot program은 5:5의 경기를 위하여 구성되어진 simulator이다. Vision 기반의 로봇축구 시스템과 같은 구성으로 되어 있으며, 입력 값은 센서에 의한 공의 좌표와, 각 로봇의 좌표 및 각도 이다. 출력은 각 로봇의 좌, 우의 바퀴의 속도를 출력하여 로봇의 움직임을 조정한다.

하지만 본 논문에서는 입력과 출력을 그대로 사용하는 것이 아니라 약간의 가공을 거쳐 사용한다. 우선 입력에서는 경기를 위하여 로봇축구 시스템을 구현할 때 가장 중요하게 판단되는 공과 공에 가까운 우리 팀 로봇의 위치, 그리고 공과 가까운 상대 팀 로봇의 위치를 선택하여 입력으로 받아들인다. 출력의 경우는 모터의 움직임을 로봇을 조정하는 것이 아니라 드리블, 킥, 이동, 수비이동 등의 일정한 기본 행동들을 호출하고 모터의 제어는 이 호출된 기본 행동들에서 직접 조정을 한다.

### 4.1 검지기

검지기는 환경으로부터 입력(센서 입력)을 2진 코드로 변환한다. 경기장을 가로 2bits 세로 2bits의 공간으로 나누어 공과 각 로봇의 좌표를 각 4bits의 코드로 나타내고 그 중에서

실제 경기의 진행에 가장 큰 영향을 미치는 공과 공에 가장 가까운 우리 팀 로봇 2대의 좌표와 공에 가장 가까운 상대 팀 로봇 1대의 좌표를 코드화 하여 총 16 bits 길이의 2진 string을 출력한다.

#### 4.2 룰 베이스

센서로부터의 입력을 분석하여 몇 개의 상태 공간으로 분류를 하고 각 상태에 따라 각각 하나의 룰 셋을 생성한다. 최초 시작을 할 때에 저장된 룰 셋이 존재할 경우 이 룰 셋을 읽어 들여 룰 베이스를 구성하고, 만일 존재하지 않을 경우에는 랜덤함수를 이용하여 룰 셋을 초기화 시켜 구성한다. 분류자 하나의 길이는 조건부 16bits와 행동부 20bits, strength 1bit, 모두 37 bits이다.

전체 입력으로부터 현재의 상태를 분석하여, 그 상태에 따라 '룰 셋 선택기'에서 적용할 룰 셋을 결정을 한다. 검지기의 입력과 선택된 룰 셋의 분류자들의 조건부를 비교하여 match set을 구성하고, 구성된 match set은 세금 (tax)을 지불한다. Match set에서 strength에 따른 roulette 선택을 하여 action set을 결정하고 'action set'에 선택된 분류자는 비용 (bid)을 지불한다. 지불된 비용은 'bucket'을 통하여 'old action set'에게 가는데, 일정의 비율(discount factor)로 감소한다. 'action set'에 결정되어진 분류자의 행동부는 효과기의 입력으로 들어간다.

#### 4.3 효과기

룰 베이스로부터 입력을 받아 키퍼로봇을 제외한 4대의 우리 팀 로봇들을 조정한다. 각 로봇의 움직임을 결정하는데 5bits의 string이 사용된다. 최초의 1bits는 로봇이 공격적인 행동을 할 것인지, 아니면 수비적인 행동을 할 것인지를 결정한다. 뒤의 4bits는 로봇이 행동할 구간을 결정하는데 사용된다. 이는 입력을 받을 때 경기장을 구분한 것과 같은 공간으로 구분하여 사용한다. 이 때 수비적인 행동을 할 경우는 배당된 공간 안에서 최대한 수비적 행동을 취하고, 공격적인 행동을 할 경우에는 다른 공격적인 행동을 하는 로봇들과 비교를 하여 공과 가까운 때에는 배정된 공간에 상관없이 공격을 하고, 그렇지 않을 경우에는 공간 안에서 가능한 수비행동을 한다.

#### 4.4 평가 및 파라미터 설정

로봇축구 시스템의 strength를 평가하기 위하여 "이벤트와 샘플링 시간"이라는 개념을 도입하였다. 이것은 일정한 시간 간격동안 행동 집합을 환경 속에서 평가하고 보상 값을 획득

하는 경우가 발생했을 때만 규칙 베이스의 갱신이 이루어 지도록 한다는 것이다[4]. 100 frame 동안을 샘플링 시간으로 보고 각 frame에서 이벤트의 발생 유무에 따라 적용된 룰의 strength를 조정하였다. 이벤트의 발생은 경기장을 두개의 골대를 기준으로 14개의 공간으로 분류(그림 2)하고, 공이 공간을 이동할 때 보상 값을 주었다. 수비지역의 평가를 F1, 공격지역의 평가를 F2를 사용하였다. 단일 룰 셋의 경우는 F1과 F2의 경우를 모두 사용하고, 클러스터링 된 룰 셋의 경우 수비 상황은 F1을 공격 상황은 F2를 사용하였다.

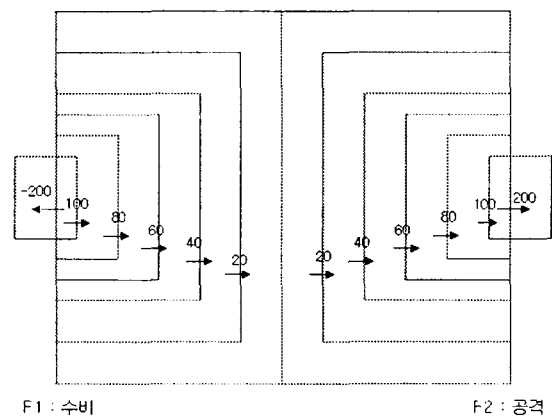


그림2. 지역에 따른 strength 평가

ZCS의 구현에 있어서 공통으로 사용된 파라미터들은 표1과 같이 설정되었다. 두 종류의 시스템에서 규칙 베이스의 population의 크기만 다르게 구성되었는데 단일 룰 셋 구조의 규칙베이스에서 population의 크기는 100이고, 공격과 수비 상황으로 클러스터링 된 룰 셋 구조의 규칙베이스에서는 각 규칙 베이스의 크기를 50으로 하여 전체적인 룰의 크기를 같게 하였다.

표 1. ZCS의 파라미터 값

Learning Rate	0.7
Tax	0
Bid	0.1
Discount factor	0.5
Crossover	0.8
Mutation	0.03

### V. Simulation 결과

전술의 변화에 따른 실시간 적응을 보기 위하여 500 sampling periods동안 하나의 전술을 이용하여 학습을 하고, 전술을 바꾸어 학습

을 하는 방법을 이용하여 3가지 전술에 대하여 연속적인 학습을 하였다.

그림3과 그림4의 strength 곡선에서 최초의 500 sampling periods는 전술 1을 이용하여 학습을 하고, 학습한 룰 셋을 이용하여 전술 2와 전술 3에 적용 및 추가학습을 한다. 두 가지 구조의 학습 속도를 보면 단일 룰 셋 구조의 분류자 시스템은 최초 400 sampling periods 정도에서 수렴하는 반면 클러스터링 된 룰 셋 구조의 분류자 시스템은 250~300 sampling periods 정도에서 수렴하여 학습 속도가 빠름을 알 수 있다. 또한 전술의 변화에 대하여 단일 룰 셋 구조는 적절한 대응을 하지 못하지만 클러스터링 된 룰 셋 구조의 분류자 시스템은 평균 strength가 변화하여 수렴함으로써 변화된 전술에 적응함을 알 수 있다.

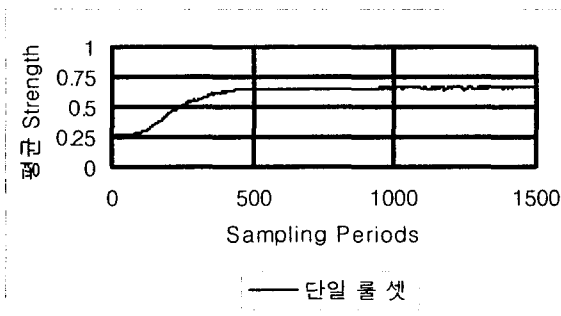


그림 3. 단일 룰 셋의 strength 곡선

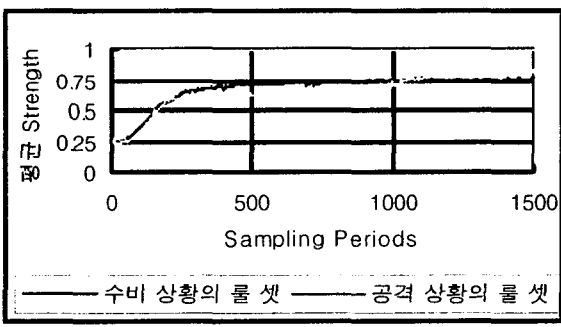


그림 4. 클러스터링 된 룰 셋의 strength 곡선

로봇 축구에서 목표로 하는 것은 결국 실점보다 많은 득점을 하여 승리하는 것이 목표이다. 두 가지 구조의 성능을 비교하기 위하여 학습을 하면서 각각의 전술에 대한 득점과 실점의 비를 표2에 나타내었다. 표2에 나타낸 값은 '득점/실점'으로 1일 경우 비기는 상황이고, 1보다 작으면 패배, 크면 승리이다. 표에 나타난 값들이 1보다 작아 모두 패했음을 알 수 있다. 하지만 클러스터링 된 룰 셋 구조의 득실점 비가 단일 룰 셋 구조의 득실점 비 보다 높은 값을 가지고 있어서 보다 좋은 성능을 가지고 있

음을 알 수 있다.

표 2. 득실점의 비(득점/실점)

	클러스터링 된 룰 셋	단일 룰 셋
전술 1	0.73	0.55
전술 2	0.46	0.32
전술 3	0.46	0.27

VI. 결 론

본 논문에서는 로봇축구와 같이 복잡한 상황을 가진 시스템에서 ZCS를 적용하였을 경우 각각의 상태에 따른 해공간이 간섭을 일으켜 유용한 규칙들이 사라지기 때문에 학습의 효율성이 떨어지는 문제점을 해결하기 위하여 상태 공간에 따라 다른 규칙 베이스를 구성하여 문제점을 해결하였다. 하지만 상태 공간을 나누는 문제에 있어서도 임의로 나누어준 공간이기 때문에 적절하게 상태가 분류되었다고는 말할 수 없다. 따라서 환경에 대한 상태공간을 분류하는데 있어서 어떤 입력을 사용하고, 어떤 분류방법을 이용하여 어떻게 나눌 것인가에 관한 연구를 통하여 시스템의 실시간 처리와 온라인 학습에 관한 성능을 극대화 시킬 것인지에 관한 연구가 추가되어야 할 것이다.

IV. 참고문헌

- [1] Gray William Flake, *The computational Beauty of Nature*, A Bradford Book The MIT Press, 1995.
- [2] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley, 1989.
- [3] Stewart W. Wilson, "ZCS: A Zeroth Level Classifier System," *Evolutionary Computation*, Vol. 2, No. 1, pp. 1-18, 1994.
- [4] 심귀보, 김지윤, "분류자 시스템을 이용한 축구 로봇의 행동전략," *한국 퍼지 및 지능시스템 학회*, Vol. 12, No. 4, pp. 289-293, 2002.
- [5] 심귀보, *인공생명의 방법론*, 드림 미디어, 2000.