

퍼지 반복 학습제어기를 이용한 동적 플랜트 제어

Fuzzy Iterative Learning Controller for Dynamic Plants

°유 학 모*, 이 연 정*

*경북대학교 전자전기공학부 (Tel: 053-950-6562; Fax: 053-950-5505; E-mail: yjlee@palgong.kyungpook.ac.kr)

Abstract In this paper, we propose a fuzzy iterative learning controller(FILC). It can control fully unknown dynamic plants through iterative learning. To design learning controllers based on the steepest descent method, it is one of the difficult problems to identify the change of plant output with respect to the change of control input($\partial e/\partial u$). To solve this problem, we propose a method as follows: first, calculate $\partial e/\partial u$ using a similarity measure and information in consecutive time steps, then adjust the fuzzy logic controller(FLC) using the sign of $\partial e/\partial u$. As learning process is iterated, the value of $\partial e/\partial u$ is reinforced. Proposed FILC has the simple architecture compared with previous other controllers. Computer simulations for an inverted pendulum system were conducted to verify the performance of the proposed FILC.

Keywords Iterative learning, Self-organizing, Steepest descent, Fuzzy logic controller, Inverted pendulum

1. 서론

지난 수 년전 부터 퍼지 학습제어기(fuzzy learning controller)에 대한 연구가 활발히 진행되어 왔다.

학습제어기란 같은 일을 수행함에 있어서 다음 수행 시에는 환경에 적응하여 제어기의 파라미터를 변화시키는 시스템을 말한다. 즉, 학습제어기는 불확실성에 대한 대처 기능(uncertainty handling capacity), 학습 기능(learning capability), 적응 기능(adaptivity), 목표 달성을 위한 최적 지향성(optimality for goal achievement) 등을 갖춘 제어기를 말한다.^{[1],[2]}

기존의 퍼지 학습제어기들은 전문가의 경험 지식이나 훈련(training) 입·출력 패턴들을 이용하여 제어기를 초기화한 후 온라인(on-line)으로 제어기를 미세조정하는 방법^{[3],[4]}과 시스템 동정화(identification)를 통해 제어기를 학습시키는 방법^[5], 그리고 시행 오차를 통하여 온라인 학습을 하는 강화 학습(reinforcement learning) 방법^{[6],[8]} 등으로 대별될 수 있다. 그러나, 앞의 첫 번째 방법은 제어기 설계시 전문가의 제어 지식과 경험에 의존도가 높고, 훈련 패턴을 얻을 수 없는 경우에는 적용하기 어려운 단점이 있으며, 두 번째 방법은 시스템 동정을 위해 전체 학습 제어 시스템이 복잡해지는 단점이 있다. 이러한 문제점을 극복하기 위해 제안된 강화 학습 방법은 설계시 설정해야 할 파라미터가 많아 실제적으로 적용하기가 용이하지 않다.

사람의 제어 행위를 살펴보면 사람은 알지 못하는 플랜트를 제어하기 위해서 몇 번이고 시행 오차과정을 거친다. 이 때, 사람은 플랜트의 파라미터 값을 학습하는 것이 아니라, 플랜트의

특성을 학습하여 다음 제어 과정시 정보로써 이용한다. 즉, 학습을 통하여 플랜트에 가해진 입력 변화와 그 때의 출력 변화 방향 정보를 획득한 후 제어에 필요한 적절한 출력을 발생시키는 것처럼 보인다. 이에 본 논문에서는 플랜트에 가해진 입력 변화와 오차 변화 방향 정보를 획득하여 제어기 출력을 생성하고 학습된 결과를 다음 제어 과정에 이용하는 이른바 퍼지 반복 학습제어기(fuzzy iterative learning controller)를 제안한다. 제안된 퍼지 반복 학습제어기는 기존 제어기에 비해 간단한 구성을 가지며 온라인 학습이 가능하다. 제안된 퍼지 반복 학습제어기의 성능 평가를 위해서 역진자 시스템 제어를 통하여 타당성을 검토한다.

본 논문은 먼저 퍼지 반복 학습제어기의 구성, 학습 방법, 플랜트 입력 변화에 대한 오차 변화 방향값 갱신에 대하여 논하고, 퍼지 반복 학습제어기의 역진자 시스템 모의 실험 결과를 고찰한다. 그리고, 마지막으로 앞으로 연구 방향에 대하여 언급하고 결론을 맺는다.

2. 퍼지 반복 학습제어기

2.1 퍼지 반복 학습제어기의 구성

본 논문에서 제안된 퍼지 반복 학습제어기 구성도는 그림 1에 서와 같다.

구성도는 퍼지 논리 제어기(fuzzy logic controller) 부분과 플랜트의 입력 변화에 대한 오차 변화 방향 값을 획득하고 학습하는 부분으로 나누어진다. 학습은 정해진 시간 동안과 제어하고자 하는 범위 내에서만 이루어진다.

퍼지 제어기(FLC)는 수학적으로 플랜트를 표현하기 복잡하거나 어려운 경우, 수학적 모델에 의존하지 않고 전문가의 경험 지

식을 통해 제어에 필요한 규칙을 생성하여 제어하는 시스템이다.

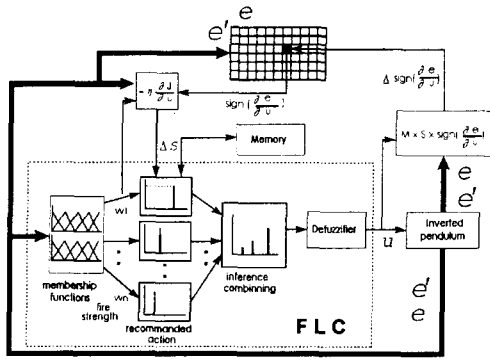


그림 1. 퍼지 반복 학습제어기 구성도
Fig. 1. Fuzzy iterative learning controller block diagram.

제어기는 퍼지화(fuzzification), 규칙 기반(rule base), 추론부(inference engine), 비퍼지화(defuzzification) 부분으로 구성되어 있다. 제어기 각 부분의 특성은 아래와 같다.

- 퍼지화: · 전건부-오차, 오차 변화분: 균일하게 분포한 삼각형 모양의 소속 함수
- 후건부-제어기 출력: 싱글톤(singleton) 소속 함수
- 규칙 기반: If-Then
- 추론부: product-sum

$$\alpha_j = \mu_{A_j}(x) \cdot \mu_{A'_j}(x') \quad (1)$$

- α_j : j번째 전건부의 적합도(fire strength)
- x, x' : 언어적 변수 A, A' : 퍼지 변수
- $\mu(\cdot)$: 소속함수

비퍼지화: 무게 중심법(center of gravity method)

식 (1)을 이용하여 FLC의 입·출력 관계는 다음과 같이 표현할 수 있다.

$$u = \frac{\sum_{j=1}^n \alpha_j w_j}{\sum_{j=1}^n \alpha_j} \quad (2)$$

여기서 n 은 규칙 총수, w_j 는 j번째 규칙의 후건부 싱글톤 위치이다.

2.2 학습 방법

제안된 퍼지 반복 학습제어기에서 제어기 출력은 다음과 같은 성능 지수를 최소화하도록 설정한다. k 번째 샘플링 순간, i 번째 반복을 나타내는 성능 지수를 다음과 같이 가정하면,

$$\begin{aligned} J^i(k) &= J_1 + J_2 \\ &= \frac{1}{2} (e^i(k))^2 + \frac{1}{2} (\dot{e}^i(k))^2 \end{aligned} \quad (3)$$

$$e^i(k) = y_d^i(k) - y^i(k) \quad \dot{e}^i(k) = e^i(k) - e^i(k-1)$$

여기서 y_d 는 원하는 출력, y 는 제어기 출력이다. 제어기 출력을 형성하는 후건부 싱글톤의 위치 변화량 ($\Delta w^i(k)$)는 최급 강하법에 의해 다음과 같이 표현될 수 있다.

$$\Delta w^i(k) = -\eta \cdot \frac{\partial J_1}{\partial w^i(k)} - \delta \cdot \frac{\partial J_2}{\partial w^i(k)} \quad (4)$$

$$= -\eta \cdot \frac{\partial J_1}{\partial e^i(k)} \cdot \frac{\partial e^i(k)}{\partial u^i(k)} \cdot \frac{\partial u^i(k)}{\partial w^i(k)} \quad (5)$$

$$= -\delta \cdot \frac{\partial J_2}{\partial e^i(k)} \cdot \frac{\partial e^i(k)}{\partial u^i(k)} \cdot \frac{\partial u^i(k)}{\partial w^i(k)} \quad (6)$$

$$= -\eta \cdot e^i(k) \cdot \frac{\partial e^i(k)}{\partial u^i(k)} \cdot \frac{\alpha_j}{\sum_{j=1}^n \alpha_j} - \delta \cdot e^i(k) \cdot \frac{\partial e^i(k)}{\partial u^i(k)} \cdot \frac{\alpha_j}{\sum_{j=1}^n \alpha_j}$$

여기서 η, δ 는 학습률(learning rate)이다.

위의 식을 이용하여 ($k+1$) 시간에 후건부 싱글톤의 위치는

$$w^i(k+1) = w^i(k) + \Delta w^i(k) \quad (7)$$

과 같이 주어진다.

위의 식에 의해 학습된 싱글톤의 위치를 메모리에 저장하여 반복 학습 시에 이용한다.

2.3 플랜트 입력 변화에 대한 오차 변화 방향값 갱신

식 (6)을 이용하여 학습할 경우 생기는 문제점은 플랜트 입력 변화에 대한 오차 변화량($\frac{\partial e}{\partial u}$)을 구하는 것이다. $\frac{\partial e}{\partial u}$ 은 같은 상태에서 제어 입력의 변화에 따른 플랜트 출력 변화를 의미하는 것으로 이 값이 플랜트 특성과 밀접하게 관계되어 있으므로, 특성을 모르는 플랜트의 학습시 이 값을 구하는 것은 어려울 뿐만 아니라 중요한 일이다. 따라서, $\frac{\partial e}{\partial u}$ 값을 구하기 위해서는 같은 상태에서 다른 입력이 가해질 경우 다음 상태의 천이를 비교해서 구해야 한다. 하지만, 일반적으로 동적 플랜트의 제어나 학습시 계속적으로 상태나 제어 입력이 바뀌므로 같은 상태에서 다음 상태의 천이를 비교하기 어렵기 때문에 $\frac{\partial e}{\partial u}$ 값을 구하는 것은 쉽지 않다. 이에 Berenji^[8]는 식 (8)과 같은 수식의 부호(sign)를 학습제어에 이용하였다.

$$\frac{\partial e}{\partial u}(k) \cong \frac{de}{du}(k) \cong \frac{e(k) - e(k-1)}{u(k) - u(k-1)} \quad (8)$$

하지만, 식 (8)의 경우 일반적으로 인접한 시간에서도 서로 다른 상태가 나타날 수 있기 때문에 이러한 근사화는 타당하지 않다. 이에 본 논문은 다음과 같은 것을 제안한다.

$\frac{\partial e}{\partial u}$ 는 상태에 따른 값으로 이를 알고 학습에 이용할 경우 모든 상태에 대한 값을 저장해야 하기 때문에 메모리 크기가 사실상 무한대가 되어야 한다. 이에 우리는 몇 개의 대표 상태를 두어 유사 척도를 이용하여 전 상태 영역을 포괄하는 방법을 제

안한다.

대표 상태(e, e')를 이용하여 상태표를 만들어 각 셀에 대하여 $\frac{\partial e}{\partial u}$ 를 임의 값으로 초기화한다. 매 시간 순간에 우리는 FILC의 출력과 플랜트의 출력 오차, 오차 변화 분을 관찰할 수 있다고 가정한다면, 다음과 같은 척도를 두어 우리가 필요로 하는 정보를 얻을 수 있다. 척도는 다음과 같다.

$$M = 1 - |e(k) - e(k-1)| \cdot |e'(k) - e'(k-1)| \quad (9)$$

$$0 \leq M \leq 1$$

척도 M 은 인접한 시간에서 e, e' 의 정보로부터 $\frac{\partial e}{\partial u}$ 의 값을 갱신하기 위해서 상태가 같음을 나타내는 척도이다. 가령, $(k-1)$ 과 (k) 시간에서 출력 결과를 각각 $y(k-1)$ 와 $y(k)$ 로 두면, 상태가 같음을 나타내는 척도 M 을 만족할 경우, 정보로서 가치를 지닌다.

이외에 출력 값이 상태표의 셀과 같음을 나타내는 척도가 필요하다. 각 셀의 소속 정도를 나타내는 척도를 S 라 하면, 이는 각 대표 상태의 소속 정도를 각 셀에 할당된 삼각형의 소속함수를 이용하여 그 소속 함수 값으로 정한다.

위와 같은 두 척도 상태가 같음을 나타내는 M 과 소속 정도를 나타내는 S 를 이용하여, 다음 식과 같이 인접한 시간에서의 출력 결과를 학습에 필요한 정보로서 이용할 수 있다.

$$\frac{\partial e}{\partial u}(k) \cong M \times S \times \text{sign}\left(\frac{e(k) - e(k-1)}{u(k) - u(k-1)}\right) \quad (10)$$

식 (10)에서 $e(k), e(k-1), e'(k), e'(k-1)$ 가 같은 상태일수록 척도 M 값은 높아져 그 상태의 $\frac{\partial e}{\partial u}$ 값 갱신을 더 많이 하게 된다. 즉, 상태가 같을수록 정보의 가치가 높아진다. 또 출력 값이 상태표의 셀과 인접할수록 척도 S 값이 높아져 $\frac{\partial e}{\partial u}$ 값 갱신량이 증가한다. 하지만, 식 (10)의 경우 인접한 상태에 한 순간에 잘못된 결과나 잡음에 의해 상태값에 대한 정보가 변하기 때문에 원하는 방향으로 제어가 되지 않을 수 있다. 이에 학습 방향이 맞을 경우 이를 강화시킬 필요가 있다. 학습 방향을 강화시키는 방법은 상태표의 상태 값 갱신을 과거의 값을 이용하여 갱신한다.

$$\frac{\partial e}{\partial u}(k) \cong M \times S \times \text{sign}\left(\frac{e(k) - e(k-1)}{u(k) - u(k-1)}\right) + \frac{\partial e}{\partial u}(k-1) \quad (11)$$

위의 식 (11)도 $\frac{\partial e}{\partial u}$ 에 대한 근사화 식이기 때문에 이 값의 부호만 학습에 이용한다.

3. 모의 실험 결과

제안된 퍼지 반복 학습제어기를 역진자 시스템(그림 3)에 적용하여 타당성을 검토하였다. 수식 및 모의 실험 파라미터는 식 (12)와 표 1과 같이 주어진다.

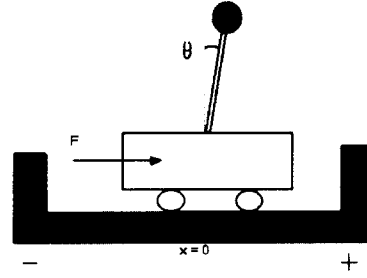


그림 3. 역진자 시스템
Fig. 3. Inverted pendulum system.

$$\ddot{\theta} = \frac{(m+M)g \sin \theta - \cos \theta (u + mL \dot{\theta}^2 \sin \theta)}{\frac{4}{3}(m+M)L - mL \cos^2 \theta} \quad (12)$$

주어진 값을 이용하여 3초($0 \leq k \leq 300$)이거나 주어진 범위(오차: $-20^\circ \sim 20^\circ$, 오차변화분: $-300^\circ/\text{sec} \sim 300^\circ/\text{sec}$)를 넘어갈 경우 한 번의 학습이 완료되어진다. 상태표는 오차, 오차 변화분에 대하여 각각 $1^\circ, 20^\circ/\text{sec}$ 으로 나누었다. 따라서, 총 상태수는 $41 \times 31 = 1271$ 이다. 반복 퍼지 학습제어기는 주어진 식 (11)에 의해 학습한 상태에 대하여 상태표의 셀 값을 갱신한다.

표 1. 모의 실험 파라미터

Table 1. Simulation parameters.

simulation parameters	
mass of the pole(m)	0.1 kg
mass of the cart(M)	1 kg
length of the pole(2L)	1 m
initial angle	10°
sampling time step	0.01 s
universe of discourse on force	-10N ~ 10N
number of membership functions	$e, e' : 5 \quad w : 25$
number of rules	25

학습률 $\eta = 2$, $\delta = 30$ 이고 w 의 초기화를 0으로 할 경우 모의 실험 결과는 그림 4와 같이 주어지고, 그림 5은 학습이 완료된 후 제어기 출력 값이다.

학습률 $\eta = 2$, $\delta = 30$ 이고, w 의 초기화를 $-5\text{N} \sim 5\text{N}$ 구간에 균일하게(uniformly random)으로 두었을 경우 제어 결과는 그림 6과 같이 주어진다. 학습률 $\eta = 1$, $\delta = 15$ 이고, w 의 초기화를 0으로 할 경우 그림 7과 같은 제어 결과를 나타낸다.

모의 실험 결과를 보면 제안한 퍼지 반복 학습제어기가 학습을 반복할수록 성능이 향상되며 제어가 잘 됨을 알 수 있다.

4. 결론

본 논문에서는 특성을 잘 모르는 동적 플랜트를 반복 학습을 통하여 원하는 성능을 갖도록 제어할 수 있는 퍼지 반복 학습제

여기를 제안하였다. 입력에 대한 오차 감소 방향 값을 찾기 위해 유사 척도를 두어 정보로서 가치를 평가하였고, 학습된 결과를 재학습시에 이용하였다. 제안된 반복 퍼지 학습제어기를 역진자 시스템에 적용하여 학습 횟수가 증가함에 따라 제어기 성능이 향상됨을 보였다.

앞으로는 학습된 결과를 원하는 방향으로 학습을 강화시키고 온라인으로 학습한 결과와 오프라인으로 학습한 결과를 서로 비교하고자 한다.

참고문헌

- [1] 변중남, "지능제어의 의미와 연구 동향," 전자공학회지, 제 22권 제 11호, pp. 1253-1262, 1996.
- [2] T. J. Procyk and E. H. Mamdani, "A Linguistic Self-Organizing Process Controller," *Automatica*, Vol. 15, pp. 15-30, 1979.
- [3] C. T. Lin and C. S. George Lee, "Neural-Network-Based Fuzzy Logic Control and Decision System," *IEEE Trans. on Computers*, Vol. 40, No. 12, pp. 1320-1336, Dec. 1991.
- [4] J. S. Roger Jang, "Self-Learning Controllers Based on Temporal Back Propagation," *IEEE Trans. on Neural Networks*, Vol. 15, No. 5, pp. 714-723, Sept. 1992.
- [5] M. H. Lee, S. Y. Lee, and C. H. Park, "Neuro-Fuzzy Identifiers and Controllers," *Journal of Intelligent and Fuzzy Systems*, Vol. 2, pp. 1-14, 1994.
- [6] A. G. Barto, R. C. Sutton, and C. W. Anderson, "Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems," *IEEE Trans. on Syst. Man. Cybern.*, Vol. 13, No. 5, pp. 834-846, 1983.
- [7] C. C. Lee, "Intelligent Control Based on Fuzzy Logic and Neural Net Theory," *Int'l Conf. on Fuzzy Logic and Neural Network*, pp. 759-764, July 1990.
- [8] H. R. Berenji and P. Khedkar, "Learning and Tuning Fuzzy Logic Controllers Through Reinforcements," *IEEE Trans. on Neural Net.* Vol. 3, No. 5, pp. 724-740, 1992.

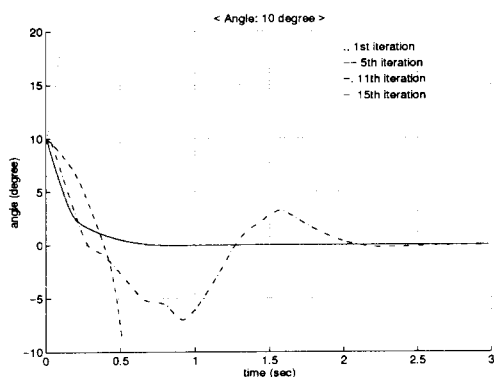


그림 4. 모의 실험결과($\eta=2$, $\delta=30$, w 초기화:0)
Fig. 4. Simulation result($\eta=2$, $\delta=30$, initial value w :0).

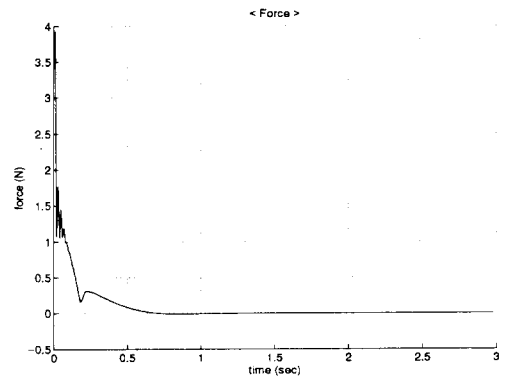


그림 5. 15회 반복 학습후 반복 퍼지 학습제어기 출력
Fig. 5. Output of the learned FILC at 15th iteration.

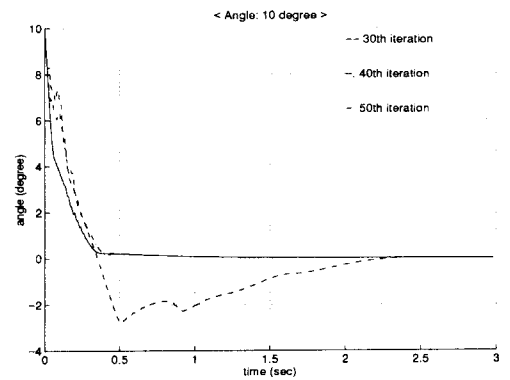


그림 6. 모의 실험 결과($\eta=2$, $\delta=30$, w 초기화: 랜덤)
Fig. 6. Simulation result($\eta=2$, $\delta=30$, initial value w : random).

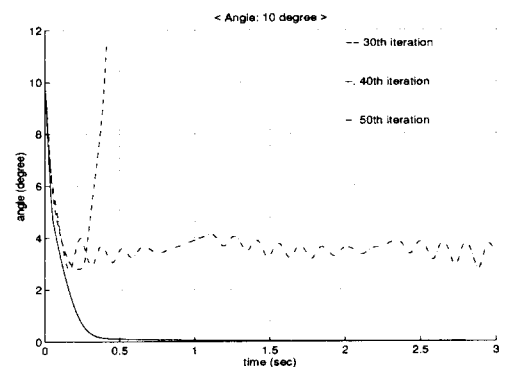


그림 7. 모의 실험 결과($\eta=1$, $\delta=15$, w 초기화:0)
Fig. 7. Simulation result($\eta=1$, $\delta=15$, initial value w :0).