

프랙탈 생산시스템에서의 강화학습을 통한 골 보정 방법¹⁾

Goal Regulation Mechanism through Reinforcement Learning in a Fractal Manufacturing System (FrMS)

신문수*, 정무영**

* 포항공과대학교 산업경영공학과 / 제품생산기술연구소 (shinms@postech.ac.kr)

** 포항공과대학교 산업경영공학과 / 제품생산기술연구소 (myjung@postech.ac.kr)

Abstract

Fractal manufacturing system (FrMS) distinguishes itself from other manufacturing systems by the fact that there is a fractal repeated at every scale. A fractal is a volatile organization which consists of goal-oriented agents referred to as AIR-units (autonomous and intelligent resource units). AIR-units unrestrictedly reconfigure fractals in accordance with their own goals. Their goals can be dynamically changed along with the environmental status. Since goals of AIR-units are represented as fuzzy models, an AIR-unit itself is a fuzzy logic controller. This paper presents a goal regulation mechanism in the FrMS. In particular, a reinforcement learning method is adopted as a regulating mechanism of the fuzzy goal model, which uses only weak reinforcement signal. Goal regulation is achieved by building a feedforward neural network to estimate compatibility level of current goals, which can then adaptively improve compatibility by using the gradient descent method. Goal-oriented features of AIR-units are also presented.

Key words: FrMS, goal-orientation, goal-regulation, Air-unit, fuzzy, reinforcement learning

1. 서론

프랙탈은 언제나 부분이 전체를 닮는 자가 유사성(self-similarity)과 순환성(recursiveness)을 갖는 형상이다(Mandelbrot 1982). 프랙탈 생산시스템(fractal manufacturing system; FrMS)은 이러한 프랙탈 고유의 특성을 리소스(resource)의 제어구조 수립에 반영하는 신개념의 생산시스템 아키텍처이다. 즉 생산시스템의 운용 과정에서 반복적으로 발견되는 임의의 패턴이 하나의 프랙탈로 정의되고, 이들의 순환적 정의를 통해 하나의 FrMS가 구축된다.

FrMS의 개념은 Ryu and Jung (2003, 2004)과 Ryu *et al.* (2003)에 의해 최초로 정립되었다. 이들의 FrMS

모델은 프랙탈의 특성을 생산시스템의 설계에 반영하고자 하는 다양한 연구(Tirpak *et al.* 1992, Warnecke 1993) 내용을 포괄하고 있다. 특히 프랙탈의 골 지향적(goal-orientated) 메커니즘과 동적 재구성(dynamic restructuring) 메커니즘을 제시함으로써 기존의 다른 시스템들과의 차별화를 실현하였다. 또한 리소스의 계층구조를 수립하고 계층적 위상에 관계없이 모든 리소스의 제어 패턴을 동일하게 설계하며, 개별 리소스의 제어 모듈을 하나의 프랙탈로 정의한다. 그러나 리소스 제어 모듈의 기능설계에 초점을 두고 있어 리소스 간의 상호 작용에 대한 고려가 부족하다.

본 논문에서 다루어지는 FrMS는 기존의 모델을 변형한 자율적이고 지능적인 리소스(autonomous and intelligent resource; AIR) 기반의 FrMS이다. 이는 프랙탈 구조가 리소스 간의 상호작용을 기반으로 설계되는 특이점을 갖는다. 본 논문은 AIR 유닛 기반 FrMS의 골 지향 메커니즘을 제안하며, AIR 유닛의 골 모델을 제시한다. 특히 퍼지추론시스템(fuzzy inference system)으로 구현되는 골의 성취도 평가 메커니즘과 강화학습(reinforcement learning) 기법을 활용한 골의 보정 메커니즘을 제시한다.

2. AIR 유닛 기반의 프랙탈 생산시스템

AIR 유닛은 실제 리소스를 의미하는 물리적인 개체인 동시에 리소스 제어를 위해 각종 정보를 처리하는 논리적인 개체이다. AIR 유닛이 의미하는 리소스는 단순한 도구적 존재가 아닌 가치창출의 주체적 존재이며, 다른 리소스들과 사용자/리소스의 고용관계를 맺는다. 또한 계층적 구조에 따라 집합적으로 정의되며, 리소스가 지닌 서비스 역량을 기반으로 표현된다. AIR 유닛은 리소스의 이러한 개념적 특성을 반영하는 자율적이고 지능적인 의사결정 개체이다. 특히 골 지향적이며, 다른 유닛들과 계약 기반의 협업을 진행한다.

AIR 유닛 기반의 FrMS는 AIR 유닛을 기본 구성 요소로 하며, AIR 유닛들의 협력 모델을 통해 프랙탈의 개념을 구현한다. 즉 계층구조상에서 AIR 유닛간 반복적으로 정의되는 사용자/리소스 관계를 기반으로 프랙탈을 정의한다. 따라서 사용자/리소스 관계를 맺는 AIR 유닛들이 연합하여 하나의 프랙탈을 형성한다.

1) 본 논문은 2004년도 한국학술진흥재단의 지원에 의하여 연구되었음 (KRF-2004-041-D00805).

AIR 유닛은 독자적인 가치모델을 바탕으로 수립되는 특정한 형태의 골을 갖는다. AIR 유닛의 골 지향성은 리소스의 역할 결정과 작업계획 수립 과정에 반영된다. AIR 유닛은 골의 달성에 보다 적합한 역할을 할당 받고, 작업계획을 수립하기 위해 능동적으로 의사결정 프로세스에 참여한다. AIR 유닛은 목표달성을 위해 능동적으로 작업을 수주하며, 효율적인 작업 수행을 위한 실행계획을 수립한다. 그림 1은 AIR 모델 기반의 프랙탈 구조의 예를 도해한 것이다.

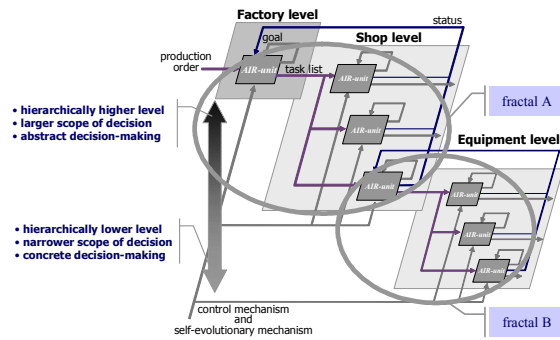


그림 1. AIR 모델 기반의 프랙탈 구조

3. Generic Goal Model (GGM)

3.1 기본 개념

생산시스템 운용의 궁극적인 목적은 가치의 창출이다. 생산시스템은 원자재에 각종 리소스의 서비스 역량을 활용한 가공을 가함으로써 제품을 생산하고 가치를 창출한다. AIR 유닛은 가치창출의 주체적 존재이며 목표지향적인 의사결정 개체이다. AIR 유닛은 독자적인 가치모델을 운용하며, 가치모델에 따라 가치창출을 극대화하기 위한 의사결정을 한다. GGM (Generic Goal Model)은 AIR 유닛의 가치모델을 정형화한다.

GGM은 AIR 유닛의 가치모델을 다음 세 가지 특성을 반영하여 수립한다.

- 이윤 (profit): 작업 수행을 통해 얻는 실제 가치와 작업 수행에 소요되는 비용의 차이
- 효용 (utility): 소속 복합체 전반의 가치창출에 공헌하는 정도
- 신용 (credit): 할당된 작업 수행에 요구되는 각종 제약조건의 이행 정도

이윤은 AIR 유닛의 의사결정 과정에서 고려되는 1차적 요소이며, 효용과 신용은 소속 복합체 전반의 이익을 위해 고려되는 2차적 요소이다. 개별 AIR 유닛이 자신의 이윤 극대화를 고집하는 경우 AIR 유닛 간의 충돌(conflict)이 발생하거나 지역최적화로 인한 각종 문제가 발생하게 된다. 효용이나 신용과 같은 2차적 요소를 함께 고려함으로써 복합체 전반의 이익을 도모하며, 장기적으로는 개별 AIR 유닛의 이윤 증대 효과도 함께 얻는다.

GGM은 위의 세 가지 특성치와 골의 성취도를 퍼지화(fuzzification)하여 표현하고, 퍼지논리규칙(fuzzy logic rule)을 정의함으로써 퍼지추론시스템을 구축한다. 특히 이를 신경망(neural network)으로 구현하며, 강화학습 기법을 기반으로 하는 골의 보정 메커니즘을 지원한다. 즉 AIR 유닛 스스로가 자신의 골을 설정하고 보정할 수 있을 뿐만 아니라 전문가의 선택적 정보

를 반영함으로써 시스템 운용의 효율을 높인다.

3.2 Fuzzy model

GGM의 개별 특성치와 골의 성취도는 다음과 같은 언어항(linguistic term)으로 표현된다.

$$\begin{aligned} Term(\text{profit}) &= \{good, bad, good, bad\} \\ Term(\text{utility}) &= \{good, bad\} \\ Term(\text{credit}) &= \{good, bad\} \\ Term(GA) &= \{good, bad, good, bad\} \end{aligned}$$

GA는 골 성취도를 의미한다. GGM의 1차적 요소인 이윤과 골 성취도는 세밀한 의사결정 지원을 위해 다른 요소들 보다 상세한 언어항을 정의한다. 개별 언어항은 삼각형(triangle) 혹은 부등변 사각형(trapezoid) 형태의 소속함수(membership function)를 갖는다. 이러한 형태의 소속함수는 미분이 불가능하지만 단순하면서도 표현력이 높다(Jouffe 1998). 개별 언어항의 소속함수는 그림 2와 같은 형태로 정의된다.

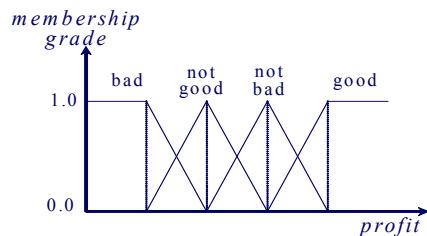


그림 2. 이윤에 대한 언어항별 소속함수

4. Goal-orientation Mechanism

AIR 유닛은 골 지향성을 갖는다. 즉 모든 의사결정에서 주어진 골의 성취를 고려하며, 골의 성취를 위한 행동계획을 수립한다. 본 장에서는 실시간 의사결정 문제를 대상으로 하는 AIR 유닛의 골 지향 의사결정 메커니즘을 제시한다. 의사결정 과정에서 AIR 유닛은 환경모델에서 발생한 임의의 이벤트에 대응하여 적절한 행동을 결정한다. 이벤트의 내용과 이벤트에 따른 반사행동들은 미리 결정되며, AIR 유닛은 여러 반사행동들 중에서 최선의 행동을 선택한다. 이 과정에서 AIR 유닛의 골 지향성이 반영된다. AIR 유닛들의 행동은 환경모델이 새로운 상태로 전이하게 한다.

AIR 유닛의 골 지향 메커니즘은 다음과 같은 구성 요소를 통해 구현된다(그림 3).

- Goal assessment network (GAN) 내부의 상태정보와 환경모델에서 발생한 이벤트 정보를 바탕으로 골의 성취도를 예측한다. 이벤트 정보는 별도의 인코딩 메커니즘에 의해 미래 상태 정보로 변환되어 입력된다.
- Goal evaluation network (GEN) 현재의 골이 환경에 적합한 정도를 평가한다. 환경에 보내는 강화 시그널을 바탕으로 평가 모델을 학습하며, 이 과정에서 골을 함께 보정한다.
- Action selection network (ASN) GAN에서 예측한 골 성취도를 바탕으로 최선의 반응행동을 결정한다. 즉 골의 성취도가 높을 것으로 예측되는 반응행동을 선택한다.
- Stochastic action modifier (SAM) 시스템의 불확실성을 고려하여 ASN이 추천하는

반응행동을 변형시킨다. 실제로 현재의 골이 시스템 전반의 운용측면에 적합하지 않을 수 있으므로 골 성취도 예측이 부정확할 수 있다.

본 논문은 AIR 유닛의 골 지향 메커니즘을 구현하는 핵심 구성요소인 GAN과 GEN을 집중적으로 다룬다.

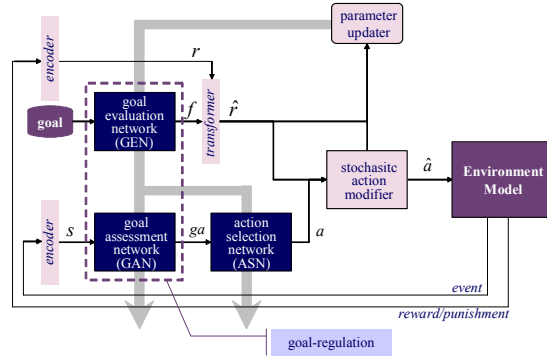


그림 3. 골 지향 메커니즘

5. Goal Assessment Network (GAN)

GAN은 임의의 상태에 대한 골의 성취도를 평가하며, GGM의 퍼지모델을 바탕으로 퍼지추론시스템을 다층 피드포워드 신경망(multi-layer feedforward network)의 형태로 구현한다. GAN은 의 구조는 그림 4와 같고, 사용되는 표기는 다음과 같다.

- $X = [x_1, x_2, \dots, x_n]$
- x_i : i 번째 입력변수
- y : 스칼라 형태의 골 성취도 수준
- T_{ij}^A : i 번째 입력변수의 j 번째 언어항
- n_i^A : i 번째 입력변수의 언어항 개수
- $T_i^A = \{T_{ij}^A | j=1, 2, \dots, n_i^A\}$
- $\mu_{A_{ij}}$: T_{ij}^A 의 소속 함수
- A_{ij} : x_i 와 $\mu_{A_{ij}}$ 로 정의되는 퍼지집합
- $A_i = \{A_{ij} | T_{ij}^A \in T_i^A\}$
- T_j^C : 출력변수의 j 번째 언어항
- n^C : 출력변수의 언어항 개수
- $T^C = \{T_j^C | j=1, 2, \dots, n^C\}$
- μ_{C_j} : T_j^C 의 소속 함수
- C_j : y_j 와 μ_{C_j} 로 정의되는 퍼지집합
- $C = \{C_j | T_j^C \in T^C\}$
- O_i^l : l 층의 i 번째 노드의 출력

GAN은 다음과 같은 퍼지논리규칙을 구현한다.

if x_1 is $A_{1l} \in A_1$, x_2 is $A_{2m} \in A_2$, and x_3 is $A_{3n} \in A_3$,
then y is $C_p \in C$

1층 입력 노드로 구성된다. 입력된 상태 정보를 다음 층으로 그대로 전달한다.

$$O_k^1 = x_k \quad \text{수식 (1)}$$

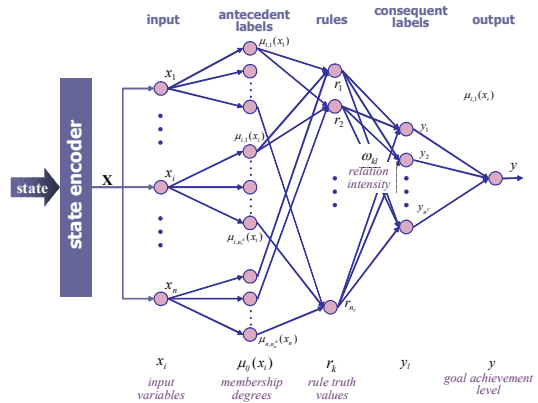


그림 4. Goal Assessment Network (GAN)

2층 전건(antecedent) 언어항 노드로 구성된다. 개별 노드는 각각 입력변수의 언어항을 상징하며, 언어항으로의 소속등급(membership grade)을 출력한다.

$$O_k^2 = \mu_{A_{ij}}(O_i^1), \quad \text{수식 (2)}$$

where $k = \sum_{i=1}^{i-1} n_i^A + j$

3층 규칙 노드로 구성된다. 개별 노드는 하나의 퍼지 논리규칙을 상징하며, 규칙의 활성화강도(firing strength)를 출력한다.

$$O_k^3 = r_k, \quad \text{수식 (3)}$$

where $r_k = \prod_{A_{ij} \in A_k^{\text{rule}}} \mu_{A_{ij}}(x_i)$

A_k^{rule} 는 k 번째 규칙에 관련한 퍼지집합들의 집합이다.

4층 후건(consequent) 언어항 노드로 구성된다. 개별 노드는 출력변수의 언어항을 상징한다.

$$O_j^4 = y_j, \quad \text{수식 (4)}$$

where $y_j = \prod_k O_k^3 \cdot w_{kj}$

w_{kj} 는 k 번째 규칙과 T_j^C 의 관련성 강도를 의미하며, l 번째 노드로부터 추정되는 골 성취도 수준은 다음과 같다.

$$\bar{y}_l = \mu_{C_l}^{-1}(y_l) \quad \text{수식 (5)}$$

5층 하나의 출력 노드로 구성된다. 출력값은 다음과 같다.

$$O_1^5 = y, \quad \text{수식 (6)}$$

where $y = f(O_1^4, O_2^4, \dots, O_n^4)$

$$= \frac{1}{n^C} \sum_l \bar{y}_l$$

6. Goal Evaluation Network (GEN)

GEN은 골의 적합도(fitness level)를 평가하는 일종의 ACE (adaptive critic element)(Barto *et al.* 1983)이다. 그림 5는 GEN의 구조를 도해한 것이다.

골의 적합도(J)는 환경에서 주어질 보상의 기대치이며, 시점 t 에서의 골이 X 일 때 골의 적합도($f_t(X)$)는 수식 (8)과 같이 정의한다.

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad \text{수식 (7)}$$

$$f_t(X) = E\{R_t | X_t = X\} \quad \text{수식 (8)}$$

γ 는 보상 할인율(discount rate)이며, 0과 1사이의 값이다. r_t 는 시점 t 에 강화 시그널(reinforcement signal)의 형태로 환경으로부터 주어지는 보상이며, 0 혹은 1의 값을 갖는다.

GEN은 GAN과 같이 전형적인 다층신경망으로 구현되며, 주어진 골에 대한 적합도($f(X)$)를 계산한다. 입력은 골을 표현하는 각종 파라미터들(x_1, x_2, \dots, x_m)이며, 한 개의 은닉층(hidden layer)을 갖는다. $f(X)$ 는 수식 (11)과 같이 계산한다.

$$v_j = \sum_{i=1}^m a_{ji} x_i \quad \text{수식 (9)}$$

$$y_j = \phi(v_j) = \frac{1}{1 + e^{-v_j}} \quad \text{수식 (10)}$$

$$f(X) = \sum_{i=1}^m b_i x_i + \sum_{j=1}^n c_j y_j \quad \text{수식 (11)}$$

$f(X)$ 는 $f_t(X)$ 의 근사값이며, 근사오차는 수식 (12)와 같다.

$$e_{t+1} = f^*(X_t) - f_t(X_t) \quad \text{수식 (12)}$$

$f^*(X)$ 는 최적의 근사치를 의미하며, Bellman(1957)의 동적계획법(dynamic programming; DP)에 근거하여 수식 (13)과 같이 정의한다.

$$f^*(X_t) \cong r_{t+1} + \gamma f_t(X_{t+1}) \quad \text{수식 (13)}$$

수식 (13)에 의해 e_{t+1} 는 수식 (14)와 같은 근사값을 갖는다.

$$\widetilde{e}_{t+1} = r_{t+1} + \gamma f_t(X_{t+1}) - f_t(X_t) \quad \text{수식 (14)}$$

AIR 유닛은 주어진 환경에 대한 최적의 골을 명시적으로 알 수는 없으므로 정확한 오차정보 또한 얻을 수 없다. 이는 Anderson(1986)의 학습모델과 유사하다. GEN은 Hamid and Khedkar (1992)와 Jouffe (1998)의 학습 모델과 같이 Sutton(1984)의 AHC(adaptive heuristic critic) 알고리즘과 역전파(back-propagation) 알고리즘(Rumelhart *et al.* 1986)에 기반하여 학습을 한다. 즉 수식 (8)과 같은 temporal difference (TD) 오차를 오차시그널(error signal)로 이용하며, gradient-descent 방식으로 파라미터를 보정한다. 파라미터 보정식은 다음과 같다.

$$\Delta c_j = \gamma \widetilde{e}_{t+1} y_j \quad \text{수식 (15)}$$

$$\Delta b_i = \gamma \widetilde{e}_{t+1} x_i \quad \text{수식 (16)}$$

$$\Delta a_{ji} = \gamma \widetilde{e}_{t+1} c_j y_j (1 - y_j) x_i \quad \text{수식 (17)}$$

$$\Delta x_i = -\gamma \sum_j \widetilde{e}_{t+1} c_j y_j (1 - y_j) a_{ji} \quad \text{수식 (18)}$$

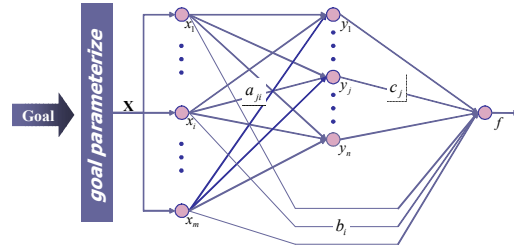


그림 5. Goal Evaluation Network (GEN)

7. 결론

본 논문은 AIR(autonomous and intelligent resource) 유닛 기반 FrMS(fractal manufacturing system)의 골 지향 메커니즘을 제시한다. 특히 AIR 유닛의 가치모델을 정형화한 GGM (generic goal model)을 제시하고, 골의 성취도 평가를 위한 GAN(goal assessment network)과 골의 평가/보정을 위한 GEN(goal evaluation network)을 제안한다. GGM은 이윤/효용/신용 3요소를 기반으로 하는 퍼지모델이며, GAN은 다층신경망으로 구현된 퍼지추론시스템이다. GEN 역시 다층신경망으로 구현되며, 강화학습 기법을 통한 학습 메커니즘을 지원한다. 본 논문은 GGM과 GAN, GEN을 바탕으로 골의 평가와 보정 알고리즘을 제시한다.

본 논문에서 제시하는 골 지향의 의사결정 메커니즘과 골 보정 메커니즘은 AIR 유닛이 스스로 판단하고 행동할 수 있는 기반을 제공한다. 특히 골 보정 메커니즘은 프랙탈의 효과적인 운용을 가능하게 한다. 하나의 프랙탈은 서로 독립적인 다양한 AIR 유닛들의 연합체로 구성되며, 환경의 변화에 따라 동적으로 변화한다. 따라서 AIR 유닛은 새로운 프랙탈에 참여하는 경우 자신의 골을 환경에 맞게 수정하여야 한다. 골이 확정적으로 주어지는 경우에는 환경의 변화에 능동적으로 대처할 수 없으며, 이는 시스템의 성능을 저하시키는 원인이 된다. 또한 환경, 즉 소속 프랙탈로부터 주어 진 강화 시그널을 바탕으로 자신의 골을 스스로 수정할 수 있으므로 시스템 설계자나 관리자의 개입 없이 시스템 스스로가 환경에 능동적으로 대처할 수 있다.

본 논문이 갖는 궁극적인 목적은 생산시스템의 자가 진화(self-evolution) 메커니즘을 개발/구현하는 것이다. 자가 진화 생산시스템은 환경의 변화에 따라 스스로 자신의 골을 형성하고 조정하며, 시스템 구조를 자유롭게 재구성한다. AIR 유닛 기반의 FrMS는 자가 진화 메커니즘을 구현하기 위한 적용 환경을 제공하며, 본 논문에서 제시되는 골 보정 메커니즘은 자가 진화 메커니즘의 핵심 요소이다. 추후 연구에서는 다음을 주제로 다룬다.

- 골 기반의 개체간 협상 메커니즘
- 골 지향의 동적 재구성 메커니즘
- 골 평가 및 보정 알고리즘의 수정/보완/검증

참고문헌

Anderson, C.W. (1986), *Learning and problem solving with multilayer connectionist systems*, Ph.D. thesis, University of Massachusetts, USA.

Barto, A.G., Sutton, R.S., and Anderson, C.W. (1983), Neuronlike adaptive elements that can solve difficult learning control problems, *IEEE Transactions on Systems, Man, and Cybernetics*, 13(5), 834-846.

Bellman, R.E. (1957), *Dynamic Programming*, NJ: Princeton Univ. Press, Princeton.

Hamid, R.B. and Khedkar, P. (1992), Learning and tuning fuzzy logic controllers through reinforcements, *IEEE Transactions on Neural Networks*, 3(5), 724-740.

Jouffe, L. (1998), Fuzzy inference system learning by reinforcement methods, *IEEE Transactions on Systems, Man, and Cybernetics -Part C*, 28(3), 338-355.

Mandelbrot, B. (1982), *The Fractal Geometry of Nature*, Freeman, New York.

Rumelhart, D., Hinton, G., and Williams, R.J. (1986), Learning representations of back-propagation errors, *Nature*, 323, 533-536.

Ryu, K. and Jung, M. (2003), Agent-based Fractal Architecture and Modelling for Developing Distributed Manufacturing Systems, *International Journal of Production Research*, 41(17), 4233-4255.

Ryu, K. and Jung, M. (2004), Goal-orientation Mechanism in the Fractal Manufacturing System, *International Journal of Production Research*, 42(11), 2207-2225.

Ryu, K., Son, Y., and Jung, M. (2003), Modeling and Specifications of Dynamic Agents in Fractal Manufacturing Systems, *Computers in Industry*, 52(2), 161-182.

Tirpak, T.M., Daniel, S.M., LaLonde, J.D., and Davis, W.J. (1992), A Note on a Fractal Architecture for Modeling and Controlling Flexible Manufacturing Systems, *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 564-567.

Sutton, R.S. (1984), *Temporal credit assignment in reinforcement learning*, Ph.D. thesis, University of Massachusetts, USA.

Warnecke, H. J. (1993), *The Fractal Company: a Revolution in Corporate Culture*, Springer-Verlag, Berlin.

부록

1. GEN의 파라미터 보정식 유도

$$E = \frac{1}{2} \sum_j y_j^2$$

$$\frac{\partial E}{\partial f} = \frac{\partial}{\partial f} \sum_j y_j = \sum_j \frac{\partial y_j}{\partial f} \cong \sum_j \tilde{y}_j$$

$$\Delta c_j = -\eta \frac{\partial E}{\partial c_j} = -\eta \frac{\partial E}{\partial f} \frac{\partial f}{\partial c_j} = \eta \tilde{y}_j$$

$$\Delta b_i = -\eta \frac{\partial E}{\partial b_i} = -\eta \frac{\partial E}{\partial f} \frac{\partial f}{\partial b_i} = \eta \tilde{x}_i$$

$$\Delta a_{ji} = -\eta \frac{\partial E}{\partial a_{ji}} = -\eta \frac{\partial E}{\partial v_j} \frac{\partial v_j}{\partial a_{ji}}$$

$$= \eta \tilde{c}_j y_j (1 - y_j) x_i$$

$$\therefore \frac{\partial E}{\partial v_j} = \frac{\partial E}{\partial f} \frac{\partial f}{\partial y_j} \frac{\partial y_j}{\partial v_j} = \tilde{c}_j y_j (1 - y_j)$$

$$\Delta x_i = -\eta \frac{\partial E}{\partial x_i}$$

$$= -\eta \sum_j \frac{\partial E}{\partial v_j} \frac{\partial v_j}{\partial x_i}$$

$$\cong -\eta \sum_j (\tilde{c}_j y_j (1 - y_j) a_{ji})$$