

Ant-Q 학습을 이용한 Gale-Shapley 문제 해결에 관한 연구

김 현[†] · 정 태 충^{††}

요 약

본 논문에서는 생물학의 개미들이 학습을 통해 목표를 획득하는 방법을 응용한 Ant-Q 알고리즘(Ant Q learning System)[1]을 Gale-Shapley[2]알고리즘을 통해 제시되었던 안정된 결혼문제(SMP: Stable Marriage Problem)[3]의 새로운 해법을 찾기 위해 적용 하였다. SMP는 남성(m_i)들과 여성(w_j)들은 각자 자신이 좋아하는 이상형에 대한 선호도(PL : preference list)를 바탕으로 안정적이면서도 최선의 짝을 찾는 것을 목표로 하고 있다. Gale-Shapley 알고리즘은 남성(혹은 여성) 위주로 안정적(stability)인 짝(Matching)을 성사시키므로 다양한 조건을 수용하지 못한다.

본 논문에 적용된 Ant-Q는 개미(Ant)의 페로몬을 활용한 학습인 ACS(Ant colony system)에 강화학습의 일종인 Q-학습[9]을 추가한 방법으로, SMP의 새로운 해법을 찾을 수 있었다.

키워드 : 안정된 결혼문제, 개미알고리즘, Q-학습, Ant-Q, TSP

Solving the Gale-Shapley Problem by Ant-Q learning

Hyun Kim[†] · TaeChoong Chung^{††}

ABSTRACT

In this paper, we propose Ant-Q learning Algorithm[1], which uses the habits of biological ants, to find a new way to solve Stable Marriage Problem(SMP)[3] presented by Gale-Shapley[2]. The issue of SMP is to find optimum matching for a stable marriage based on their preference lists (PL). The problem of Gale-Shapley algorithm is to get a stable matching for only male (or female). We propose other way to satisfy various requirements for SMP.

ACS(Ant colony system) is an swarm intelligence method to find optimal solution by using pheromone of ants. We try to improve ACS technique by adding Q learning[9] concept. This Ant-Q method can solve SMP problem for various requirements. The experiment results shows the proposed method is good for the problem.

Keywords : Stable Marriage Problem, Ant Colony System, Q-learning, Ant-Q, TSP

1. 서 론

인간은 본능적으로 종족번식을 위해 자신이 원하는 이성을 찾기 위해 끊임없이 탐색하며 노력을 기울인다. 짧은 시간, 혹은 긴 시간동안 만남과 헤어짐을 반복하여 결국 이성을 만나게 되며 결혼을 하게 된다. 본 논문에서는 이러한 인간의 본능적인 욕구를 충족할 수 있는 매칭에 관련된 방법을 찾기 위해 Gale-Shapley[2]연구를 통해 잘 알려진

SMP를 기반으로 하는 연구를 진행하였다. SMP는 안정된 결혼이라는 목적을 가지고 n 명의 남자와 n 명의 여자들이 서로가 원하는 이상형의 우선순위(PL)를 정하고 되도록 자신의 이상형과의 결혼을 하기 위한 매칭 방법들을 찾기 위해 연구되어 왔다[3,4,5,6,7].

SMP는 기숙사 배정문제(Roommates Problem)[4]와 같은 다양한 조건에서 매칭을 위한 새로운 문제를 만들었다. 기본 SMP의 해결법은 매칭을 위해서는 지켜야 할 필수 규칙(선택의 시작은 남자가 먼저 시작하며 선택한 여자가 남자를 싫어할 경우 선택의 우선권이 다른 남자에게 넘어간다)이 있다.[6] 남녀는 각각 (그림 1)처럼 자신이 좋아하는 이상형에 대해 이상형의 리스트(PL : preference list)를 만들 수 있고 작성된 PL 안의 우선순위를 바탕으로 상대방의 PL 과

※ 이 논문은 2008년도 경희대학교 지원에 의한 결과임.

† 정 회 원 : 경희대학교 컴퓨터공학과 박사

†† 정 회 원 : 경희대학교 컴퓨터공학과 교수

논문접수: 2010년 8월 4일

수정일: 1차 2010년 10월 14일, 2차 2010년 11월 1일

심사완료: 2010년 11월 29일

비교를 통해 이상형과 만나게 된다. 예를 들어 남자를 m_i 여자를 w_j 라 한다면, 서로는 PL 을 만들고 규칙에 의해 남자인 m_i 가 먼저 마음에 드는 w_j 를 선택하게 된다. 이때 선택된 w_j 는 자신을 선택한 m_i 가 PL 에서 우선순위가 가장 높다면 매칭을 이루는 것이고 가 w_j 의 PL 에서 우선순위가 낮을 경우에는 이루어지지 않고 두 번째 m_i 에게 선택의 기회가 돌아가게 된다. 이런 규칙을 따라 m_i 와 w_j 는 자신이 좋아하는 이상형을 선택할 수 있게 되는 것이다. 하지만 SMP가 가지고 있는 문제점은 선택의 우선권은 오직 m_i 에게 주어진다라는 것이며 우선순위가 높은 위치의 m_i 에게 우선권이 주어지게 된다는 것이다. 이것은 어느 한 쪽에만 유리한 해법이라 할 수 있겠다. 만약, 이전에 대부분의 m_i, w_j 가 매칭을 이루었을 때 후순 순위에 매칭을 하게 되는 m_i 와 w_j 의 경우에는 자칫 최악의 매칭을 이룰 수도 있다는 문제점을 가지고 있었다[2,3].

본 논문에서는 모든 m_i 와 w_j 의 PL 을 고려한 평등하고 안정된 매칭을 이룰 수 있는 해법을 제시하고자 하였으며 이것은 남녀평등을 지향하며, 처음 선택자와 마지막 선택자에게도 공평하게 선택의 기회를 주어 모든 선택자가 자신의 이상형을 안정적으로 만날 수 있도록 Ant-Q알고리즘을 적용하였다. Ant-Q는 기존의 생물학에서 연구되어 진 개미집단의 습성을 기초로 하는 ACS에 정확성을 더하기 위해 Q-학습이라는 요소가 추가된 알고리즘이다[9].

본 논문에서는 Ant-Q의 기본 아이디어를 SMP에 적용하였다. 이것은 개미가 먹이를 찾아 목적지까지 탐험하는 과정을 m_i 가 PL 을 기반으로 w_j 를 선택하는 과정에 적용하였다. 개미는 임의의 남자 m_i 로부터 출발하는 에이전트(Agent)이며 m_i 의 PL 에서 우선순위를 기준으로 w_j 까지 이동을 시작한다. 남자 m_i 로부터 출발한 개미는 Ant-Q 알고리즘에서 선택의 가장 중요한 평가 요소인 페로몬을 분비하며 m_i 의 PL 안에 우선순위 w_j 로 이동하게 된다. 이 행동은 각각의 m_i 가 각각의 w_j 를 선택할 때 반복 된다. 마지막 m_i 까지 자신이 좋아하는 이상형을 선택할 때 Ant-Q의 알고리즘이 적용되어진다.

논문은 다음과 같이 구성된다. 본 장인 서론에 이어 2장에서는 SMP, Ant-Q와 관련 연구들을 살펴보고, 3장에서는 Gale-Shapley 알고리즘으로 해결했던 SMP를 Ant-Q알고리즘을 적용한 새로운 해법을 제시하며, 4장에서 해법을 적용하여 얻게된 실험결과와 알고리즘의 성능, 그리고 본 논문의 결론을 이야기 하고자 한다.

2. 관련 연구

2.1 SMP

본 논문에서 연구되는 SMP는 원래 1962년에 Gale와

Shapley가 안정적인 짝을 찾는 문제에 관련된 논문을 처음으로 발표하였다. 논문에서 소개된 안정적 짝을 이루는 문제는 SMP(Stable marriage problem)라는 이름으로 의사실습생을 병원에 배치하는 알고리즘으로 10년간 미국 전역의 병원에서 사용되어 졌다[15].

SMP는 이 세상에 같은 수 n 명의 남자와 여자가 있다는 가정으로 서로가 좋아하는 이상형의 PL 을 만들고 이것을 바탕으로 안정된 매칭을 이루는 것을 그 목표로 하고 있다. SMP에 정의하고 있는 Gale-Shapley의 해법은 <표 1>과 같은 구성원들의 PL 을 기준으로 모든 m_i, w_j 가 각각 안정된 매칭(Stable matching)을 이룰 때 얻을 수 있다. 여기서 정의되는 안정성(Stable)이란 (m_i, w_j) 의 형태로 커플이 되면 더 이상 변동이 없어야 함을 의미하며, (m_i, w_j) 혹은 (w_j, m_i) 형태의 매칭 결과를 얻을 수 있다. 이것은 수식(1)과 같이 정의할 수 있다. $mr(m,w)$ 은 남자의 PL 안에서 W (여자)의 순서이며, $wr(m,w)$ 은 여자의 PL 안에서의 M (남자)순서를 의미한다.

$$sm(M_i) = \sum_{(m,w) \in M_i} mr(m,w) \quad (1)$$

$$sw(M_i) = \sum_{(m,w) \in M_i} wr(m,w)$$

본 논문에서는 m_i 를 중심으로 한 선택 결과를 정의할 때 $sm(M_i)$ -(남자가 여자를 선택), w_j 를 중심으로 할 때는 $sw(M_i)$ -(여자가 남자를 선택)와 같은 기존의 SMP의 정의를 기반으로 연계 되는 결과 이외에 새로운 정의를 통한 확장된 결과를 얻고자 하였다.

정의(1)은 수식(1)을 바탕으로 SMP의 새로운 해법을 위한 정의이다. 이것은 PL 에서 이상형은 가장 우선순위에 있으며 그렇지 않은 경우에서 후순위로 밀리는 것에 착안하여 정의되었다.

정의1 $sw(M_i)$ 가 최대이고 $M_i \in M$ 을 만족할 때 남자를 최적으로 하는 안정된 매칭이 성립된다.

정의2 $sm(M_i)$ 가 최대이고 $M_i \in M$ 을 만족할 때 여자를 최적으로 하는 안정된 매칭이 성립된다.

정의3 $(sm(M_i) + sw(M_i))$ 가 최소값을 가지고 $M_i \in M$ 을 만족할 때 모두 구성원 PL 을 고려한 평등한 매칭을 이룰 수 있다.

정의4 $|sm(M_i) - sw(M_i)|$ 가 최소값을 가지고 $M_i \in M$ 을 만족할 때 남녀 각각의 PL 을 모두 고려하여 서로 공평하게 매칭을 이룰 수 있다.

(정의 1) 다양한 조건을 고려한 안정된 매칭

2.2 Ant-Q

개미들은 에이전트(Agent)이며 등지로 부터 음식물을 찾아 탐험하고 발견된 음식물을 자신의 등지로 이동시킨다. 이때 탐험과정 동안 자신이 이동한 경로에 대해 페로몬(pheromone)을 분비하고, 이 후 다른 다수의 개미들은 분비된 페로몬을 따라 이동 후에 나머지 먹이를 등지로 옮겨오는 역할을 한다. 이때 페로몬은 시간이 지날수록 산화성으로 인해 산도(Acidity)가 약해져 많은 개미들, 혹은 초기에 이동한 개미가 먹이를 찾은 후 등지로 돌아올 때 이를 감안해 새롭게 페로몬을 분비하여 갱신과정(update)을 거친다. 새롭게 갱신된 페로몬은 기존의 페로몬에 추가되어 산도가 더욱 높아지며, 결국 산도가 약한 다른 경로는 무시되고 높은 산도를 가진 경로가 최종 이동경로로 결정된다. 이러한 개미(Agent)의 이동과정을 알고리즘으로 표현한 것을 AS(Ant system), 그리고 이것의 효율성을 높인것을 ACS(Ant colony system)라 한다[1,8,10,11].

Ant-Q는 기존의 AS에 Q-학습(Q-learning)을 적용시킨 것으로 기존의 알고리즘들의 장점을 결합하여 만들어 졌다. Q-학습(Q-learning)은 C.J.C.H Watkins가 제안한 강화학습의 단점을 보완하는 학습 방법으로 통계적 동적 프로그래밍(stochastic dynamic programming)에 근거한 학습방법이다. Ant-Q 모델은 상태공간을 하나의 에이전트로 탐색하는 일반적인 Q-학습과는 달리 협력 에이전트들의 집합을 이용해 학습을 수행하게 된다. 이러한 에이전트들은 서로 협력하여 'AQ-Value'으로 표현되는 정보를 교환하게 된다. 만는 목적지 까지 지정된 경로를 따라 이동하게 되며 이때 발생한 페로몬의 양을 바탕으로 학습을 하게 된다. Q-학습[9]의 관점에서 재해석된 강화학습방법이지만, Ant-Q 모델은 상태공간을 하나의 에이전트로 탐색하는 일반적인 Q-학습과는 달리 협력 에이전트들의 집합을 이용해 학습을 수행하게 된다. 이러한 에이전트들은 서로 협력하여 AQ-Value(Q-학습에서 Q-값과 유사)으로 표현되는 정보를 교환한다[1].

Ant-Q에서 노드(r)에 있는 에이전트 k 가 노드(u)로의 이동은 수식(2)와 같이 상태전이 규칙(state transition rule)에 의해 수행된다.

$$s = \begin{cases} \arg \max_{u \in J_k(r)} [AQ(r,u)]^q \cdot HE(r,u)^\beta & \text{if } q \leq q_0 \text{ (exploitation)} \\ S & \text{otherwise (exploitation)} \end{cases} \quad (2)$$

$AQ(r,u)$ 는 Ant-Q값으로, m_i 의 PL 에서 w_j 까지의 거리인 간선 $E(r,u)$ 에 관계된 양의 값(positive value)이다. $AQ(r,u)$ 는 Q-학습의 Q값과 상응하는 값으로, 본 논문에서는 노드(m_i 와 w_j 매칭되었을 때 m_i 의 PL 에서의 w_j 까지의 거리 : r)에서 노드(다음 매칭에서 생성되는 m_i 와 w_j PL 거리 : u)로 이동하는 것이 얼마나 유용한지를 나타내는 것으로 이동함에 따라 그 값이 갱신된다.

$AQ(i,j) = AQ_0 = 1 / (\text{average length of edges.}n)$ 이 다 .

$HE(r,u)$ 은 현재 노드(m_i)에서 노드(w_j)를 선택할 때 우수성을 평가하는 휴리스틱 값이다. $J_k(r)$ 은 현재 노드(m_i)에서 에이전트 k 가 방문할 수 있는 남아 있는 노드들(남아있는 w_j)의 집합이며, 파라미터 δ 와 β 는 AQ-Value와 휴리스틱 값의 상대적 중요도를 나타낸다. q 는 [0,1]사이에서 정규적으로 분포된 무작위 파라미터이고, q_0 는 [0,1]사이의 값을 가지는 파라미터를 나타낸다. S 는 m_i 가 w_j 를 선택할 때 수식(3)처럼 표현 할 수 있으며 이것은 확률분포에 따라서 선택된 임의의 변수이다.

$$p_k(i,j) = \begin{cases} \frac{[AQ(i,j) \cdot [HE(i,j)]^\beta]}{\sum_{u \in J_k(r)} [AQ(i,j) \cdot [HE(i,j)]^\beta]} & \text{if } s \in J_k(r) \\ 0 & \end{cases} \quad (3)$$

Ant-Q의 목표는 기존의 AS에서 문제시 되던 비슷한 여러 가지의 해에서 확률적으로 보다 나은 해 값을 찾을 수 있는 AQ값을 찾는 것으로 이것은 Q-학습을 기반으로 수식(4)에 의해 갱신과정을 거쳐 생성된다.

$$AQ(i,j) \leftarrow (1-\alpha) \cdot AQ(i,j) + \alpha \cdot (\Delta AQ(i,j) + \gamma \cdot \text{Max}AQ(s,z)) \quad (4)$$

α ($0 < \alpha < 1$)는 페로몬 지연 파라미터로 학습율(learning rate)이며, γ 는 할인율(discount rate)이다.

3. Ant-Q를 적용한 SMP의 새로운 해법

본 논문에서는 Ant-Q를 SMP에 적용하였고, 정의 3을 통해 본 논문에 응용한 Ant-Q를 정의 2에서는 Ant-Q를 적용한 새로운 SMP의 알고리즘을 명시하였다. 임의의 남자(m_i)로부터 출발하는 각각의 개미는 m_i 의 PL 의 우선순위 w_j 로 이동을 하며 페로몬을 분비하게 된다. 이때 페로몬의 상태 값(분비된 양)이 가장 큰 경로의 끝에 위치한 w_j 가 m_i 와 매칭을 이루게 된다. 그리고 또 다른 임의의 m_i 로부터 출발한 개미는 PL 의 순서대로 w_j 를 선택하게 되며 다음 차례 m_i 의 순서대로 이동을 하게 되며 모든 m_i 가 같은 과정을 수행한다. 이때 에이전트가 생성한 페로몬은 수식(3)을 기반으로 하는 상태 변화법칙에 따라 가장 많은 페로몬이 분비된 m_i, w_j 사이의 경로(PL 에서 최대한 우선순위)가 최적화된 경로라 할 수 있겠다. 여기서 페로몬갱신법칙 [1]은 수식(2)로 나타내며 개미들이 자주 방문하는 곳에 페로몬의 양을 증가시켜 매칭에 도움을 주기 위해 적용되었다. 에이전트의 이동법칙은 Ant-Q의 기본 알고리즘에 의해 개미 k 가 i 번째 남자가 j 번째 여자를 선택할 확률 값을 가지는데 사용된다.

Stable marriage problem(SMP)

SMP는 남녀 각각의 이상형을 찾아 안정된 매칭을 이루는 것을 해결하는 것이 문제의 핵심이다.

수식(2)(3)를 기반으로 $m_i(i=\{1,\dots,n\}),w_j(j=\{1,\dots,n\})$ 는 매칭을 위한 남자, 여자 목록이며, $A = \{(i,j) : i \in M, j \in W\}$ 는 매칭이 된 (m_i, w_j) 를 의미한다. Q 는 수식(4)를 기반으로 최종 매칭을 결정짓는 값이다. $\delta(i,j), \tau(i,j)$ 는 m_i, w_j 사이의 거리와 개미로부터 분비된 페로몬의 크기값을 결정 짓는 것으로 매칭 과정에서 해가 한 가지 이상일 때 최적의 매칭을 결정짓는 역할을 한다.

단, $\delta(i,j) \neq \tau(i,j)$ 이면 SMP는 문제의 해법을 찾지 못한다.

(정의 2) Ant-Q가 적용된 SMP 알고리즘

```

Initialize
for 모든 남자  $m$ 에 개미가 생성된다.

Loop /* at this level each loop is called a step */
    모든 개미는 페로몬을 분비하며  $m$ 의 PL에 있는  $W$ 로 이동한다.

Until
    페로몬 업데이트 룰에 의해 가장 큰 페로몬값이 형성되어 매칭이 이루어질 때까지 계속된다.
    AQ++ / AQ값이 계속 증가한다.
Until End_condition
    
```

(정의 3) Ant-Q 알고리즘

본 논문에서는 새롭게 정의된 Ant-Q알고리즘의 SMP에서의 성능 실험을 위해 m_i, w_j 의 크기를 ACS를 적용했던 기존의 논문[12]과 같은 환경에서 진행하였다.

에이전트의 상태 이동 범칙은 i 번째 m 에서 출발한 개미는 수식(2)의 룰에 따라 w_j 를 선택하게 된다. s 는 수식(3)의 확률분포(Probability distribution)에 의해 따라 선택된 무작위 파라미터이다. 이것은 m 으로부터 출발한 개미가 PL 안의 w 를 선택하였을 때 발생하는 페로몬이며 이는 간선 길이의 연산만으로 다음 노드를 선택하는 기존의 ACS[14]와 달리, 확률 분포를 이용해서 다음 노드를 선택하는 과정이 추가됨으로써 ACS가 가지고 있는 국부 최적에 빠지기 쉬운 한계에서 어느 정도 벗어나고자 하고 있다. 특

히 개미 k 가 노드(i)에서 노드(j)로 이동할 확률을 계산하는데, $q < q_0$ 인 경우 에이전트는 탐색 행동(exploitation action)을 취하게 된다. 이것은 장시간동안 축적된 정보(long-term)를 최선의 선택으로 하여 분비된 페로몬 정보와 단시간 생성된 정보(short-term)로 이동 거리와 관련된 휴리스틱 값을 사용하게 되며 확률 값은 페로몬의 양이 많고 이동한 간선의 거리 값이 작을수록 큰 값이 되어 최적이 된다. 반대로 $q > q_0$ 인 경우에는 에이전트는 편향된 탐험을 수행하기 위해 확률 분포(S)를 사용하여 길이가 짧고 많은 양의 페로몬을 가진 간선 선택을 선호하게 된다. 이때 확률 값이 큰 노드가 선택될 가능성이 높다.

본 논문에서는 Ant-Q를 바탕으로, 수식(2)에 따라 개미는 모든 남녀 각각의 이상형에 대한 짝을 찾아 이동할 때 이동경로가 만들어지고 페로몬을 분비하며 이를 갱신하는데 여기서 발생하는 페로몬의 갱신식을 수식(4)로 정의 할 수 있었다.

여기서,

$$\Delta AQ(i,j) = \begin{cases} F & \text{if } (i,j) \in \text{global best matching done by ant } k \\ 0 & \text{otherwise} \end{cases} \quad \text{의}$$

조건을 따라서 PL 의 우선순위 이상형을 만나 안정된 매칭을 이룬 $(m_i, w_j) M$ 의 함수 F 는 다음과 정의한다.

$$F = f(M) + AQ \quad (5)$$

$f(M)$: 안정된 매칭을 이룬 남녀의 수($0 \leq f(M) \leq n$)

AQ : 안정된 매칭을 결정하는데 참조하게 되는 변수 값이다.

수식(4)에서 α ($0 < \alpha < 1$)는 페로몬의 증발을 억제하여 데이터 손실을 막고자하는 지연 파라미터이고(시간이 지날수록 페로몬이 증발한다), $\Delta \tau(i,j)$ 는 초기에 분비된 페로몬의 양이다. 여기서 F 는 개미 k 가 PL 에서의 모든 대상으로 이동 했을 때 발생된 전체 경로에서 찾게된 최적의 경로의 길. 즉, 개미 k 가 남녀의 PL 에서 우선순위인 서로의 이상형으로 이동하여 매칭을 이루게 되는 것을 의미한다.

$m_i(i=10)$	$w_j(j=10)$
$m_1 = \{10,1,9,2,4,5,3,6,8,7\}$	$w_1 = \{9,4,7,10,3,8,1,5,2,6\}$
$m_2 = \{1,8,10,2,4,5,9,3,7,6\}$	$w_2 = \{5,3,9,4,2,10,1,8,6,7\}$
$m_3 = \{5,8,1,4,2,3,6,7,9,10\}$	$w_3 = \{10,6,8,9,2,4,3,7,5,1\}$
$m_4 = \{9,8,10,3,2,4,5,6,1,7\}$	$w_4 = \{9,5,10,6,8,3,4,7,1,2\}$
$m_5 = \{9,6,5,4,10,8,1,7,2,3\}$	$w_5 = \{5,8,1,4,2,3,6,7,9,10\}$
$m_6 = \{7,4,10,2,5,6,8,1,9,3\}$	$w_6 = \{9,8,10,3,2,4,5,6,1,7\}$
$m_7 = \{1,10,5,6,3,7,2,9,4,8\}$	$w_7 = \{1,3,9,5,2,8,6,4,10,7\}$
$m_8 = \{4,2,9,7,5,6,8,1,10,3\}$	$w_8 = \{8,6,2,9,5,1,7,4,10,3\}$
$m_9 = \{9,8,5,6,3,7,2,1,4,10\}$	$w_9 = \{2,5,10,8,3,6,9,4,7,1\}$
$m_{10} = \{4,7,1,3,9,5,8,2,10,6\}$	$w_{10} = \{7,9,5,4,1,10,6,2,8,3\}$

(그림 1) 모든 m_i, w_j 의 PL (preference list)

(그림 1) (m_i :남자의 PL , w_j : 여자의 PL)은 본 논문에서 제시한 알고리즘의 평가를 위한 실험을 위해 m_i, w_j 의 크기를 각각 10으로 생성하였다.

$AQ(i, j)$ 는 매칭을 하기 위한 (m_i, w_j) 사이 즉 노드 (i)과 노드(j)사이의 페로몬 양으로, 전체노드에서 선정된 노드가 최적 경로, 즉 서로가 원하는 이상형이면 F 로 설정하고 그렇지 않으면 0으로 설정된다. 이것은 페로몬의 갱신을 위해서 증가된 값의 반영으로 기존에 분비된 페로몬의 합에 비중을 두고 할당한다. 페로몬 갱신은 개미가 방문한 남녀들의 새로운 페로몬의 증가 값과 기존 페로몬의 증발량을 고려하여 최종 페로몬의 총합을 구하면 완성 된다. (그림 2)는 (그림 1)을 기반으로 수식(5)를 적용한 매칭의 결과들이다.

M_1	(1,10),(2,8),(3,5),(4,3),(5,6),(6,7),(7,1),(8,2),(9,9),(10,4)
M_2	(1,7),(2,9),(3,2),(4,1),(5,5),(6,3),(7,10),(8,8),(9,6),(10,4)
M_3	(1,10),(2,8),(3,5),(4,2),(5,9),(6,7),(7,1),(8,4),(9,6),(10,3)
M_4	(1,6),(2,5),(3,3),(4,10),(5,9),(6,4),(7,1),(8,7),(9,8),(10,2)

(그림 2) 안정된 매칭의 결과들

<표 1>은 (정의1)을 기반으로 (그림 12)에서의 m_i, w_j 의 매칭을 계산한 4가지 형태의 새로운 안정된 매칭의 결과 값을 의미한다. 이것은 다음과 같이 설명할 수 있다. 그림 2의 (1,7)은 w_7 의 PL 안에 m_1 이 가장 우선순위에 있는데 이때 에이전트가 이동하게 되는 PL 에서의 w_5 의 거리 값은 '1'이다. 즉 m_1 으로 부터의 거리 값이 가까울수록 이상형에 가깝다는 것이다. 멀어질수록 그 반대라 하겠다. 이런 방법으로 개미 k 는 m_i, w_j 의 거리를 계산하게 되는데 m_1 으로 부터 되도록 가까이에 위치한 w_j 들의 거리 값을 더해 그 값이 가장 작은 것이 최적의 해가 되며 이것을 <표 1>과 같이 만들 수 있다. sm 은 남자인 m_i 를 기준으로 이루어진 안정된 매칭의 결과이며, sw 는 여자인 w_j 를 기준으로 이루어진 안정된 매칭의 결과(횟수)이다. $sm + sw$ 는 m_i, w_j 모든 구성원의 PL 을 반영해 만든 안정된 평등 매칭의 결과이고, $|sm - sw|$ 는 m_i 와 w_j 어느 한쪽에 치우치지 않고 서로를 선택하여 이루어진 안정된 매칭을 이룬 결과를 의미한다. M_1 은 남자가 여자를 선택하여 이루어진 매칭의 경우의 수이고 M_2 는 반대로 여자가 남자를 선택하여 매칭을 이룬 결과를 보여주고 있다. M_3 의 $sm + sw$ 가 '58'로 이것은 (정의1)에 의거해 모든 구성원을 만족하는 매칭을 이룬 결과이다. (정의1)에서 언급한 정의4인 $|sm - sw|$ 의 경우엔 (그림 2)와 같이 M_4 의 결과에서 가장 최적의 해는 '1'이 되는 것이다.

이것은 기존의 SMP의 해법을 벗어나 본 논문에서 찾고자 하는 SMP의 새로운 해법이라 할 수 있겠다.

다음은 수식(5)를 완성하기 위한 AQ 를 구하는 방법으로 (정의1)과 (정의3)에 의해 생성할 수 있다.

3.1 남자의 PL 을 우선으로 고려한 안정된 매칭

(정의1-1)을 따라서 만일 안정된 만남을 이룬 M 이 sw 의 최대 값을 가지고 있으면 남자의 입장에서의 최적화된 결과라 할 수 있다. $\eta(i, j) = w(i, j)$ 라 했을 때 수식(3)의 페로몬을 (i, j)에 곱하여 값 $\eta(i, j)$ 을 구한다. 남자가 우선시되는 안정된 매칭을 결정하는 변수 값 AQ 는 수식(6)과 같이 정의할 수 있다.

$$AQ = \frac{\prod_{i=1}^{n-1} (f(M) - i)}{(n-1)!} sw \quad (6)$$

여기서 sw 는 w_j 의 PL 에서의 m_i 의 순위를 크기값으로 나타낸것의 총합을 의미한다.

$f(M) = n$ (M 은 안정된 매칭)이고 $AQ = sw$ 이면, 수식(5)에 따라 $f(M) = n + sw$ 가 성립된다.

수식(6)을 통해 AQ 의 최대 값을 구하여, m_i 를 우선으로 하는 안정된 매칭을 이룰 수 있다.

3.2 여자의 PL 을 우선으로 고려한 안정된 매칭

(정의1-2)를 따라, 만일 안정된 매칭 M 이 sm 의 최대 값을 가지면 w_j 를 위주로 하는 안정된 매칭을 이룰 수 있다. ($n(i, j) = mr(i, j)$ 일 때 i 번째 m_i 의 PL 에 의거한 j 번째 여자의 등급) 이 방법으로 w_j 의 PL 에 의거해 여자의 높은 등급과 페로몬의 큰 합을 가진 쌍들을 선택한다.

여기서 AQ 는 다음과 같이 수식(6)의 sw 를 sm 으로 바꾼다.

$$AQ = \frac{\prod_{i=1}^{n-1} (f(M) - i)}{(n-1)!} sm \quad (7)$$

이고 $AQ = sm$ 이라면, $F = n + sm$ 이다.

(정의1-2)에서 만일, sm 이 최대값을 가지고 $f(M) = nM$ 이 안정된 매칭이고, 그 해법이 w_j 최적의 안정된 매칭 이라고 할 때, w_j 를 우선으로 하는 안정된 매칭을 구할 수 있다.

3.3 모든 사람의 PL 을 고려한 평등적 안정된 매칭.

(정의1-3)에서 만일, 안정된 매칭 M 이 ($sm + sw$)을 가지면 이것은 모든 m_i, w_j 의 PL 을 고려하여 안정된 매칭을 이룬 것이라 할 수 있겠다. 수식(2)(3)에서

$HE(i, j) = (mr(i, j) + wr(i, j))^{-1}$ 는, i 번째 남자 PL 에서 j 번째 여자와 j 번째 여자 PL 의 i 번째 남자의 합의 역수를 계산한다. 이것은 순위의 합이 작을수록 $HE(i, j)$ 가 커진다.

AQ 는 수식(6)(7)과는 조금 다르다.

$$AQ = \frac{\prod_{i=1}^{n-1} (f(M) - i)}{(n-1)!} \cdot \frac{1}{(sm + sw)} \quad (8)$$

$F = f(M)$ 이고 $f(M) < n$ ($AQ=0$ 이면, 모든 $f(M) < n$ 이기 때문에)이다. 그리고 $AQ = sm$ 이고 $f(M) = n$ (만약 M 이 안정된 매칭이면)이면 $F = n + sm$ 이 된다.

그리고 (정의1-3)처럼 $sm + sw$ 가 최소값을 가지고 $f(M) = n$ 이면, $F = n + (sw + sm)^{-1}$ 이 된다. 이 해법은 모든 남녀의 PL 을 고려한 평등을 추구하는 안정된 매칭을 만들 수 있다. 함수 F 를 최소화 할때, 모든 남녀가 평등한 안정된 매칭을 구할 수 있는 것이다.

3.4 남녀에게 공평한 안정된 매칭

(정의1-4)에 의거해, 안정된 매칭 M 이 $|sm - sw|$ 에 의해 최소 값을 갖는다면 m_i, w_j 를 각각 고려한 안정된 매칭을 찾는 것이며, 이것은 어느 한쪽의 PL 만을 우선시 하지 않고 남녀 각각의 PL 을 모두 고려하게 된다. 이것은

$(0 \leq mr(i, j) - wr(i, j) \leq n - 1)$ 의 범위에서 다음 수식을 따라 값을 얻을 수 있다.

$$HE(i, j) = \begin{cases} |mr(i, j) - wr(i, j)|^{-1} & mr(i, j) \neq wr(i, j) \\ 2 & otherwise \end{cases}$$

그 누구에게도 우선권이 없이 남녀 모두에게 공평하게 PL 에 있는 이상형을 찾기 위해 Ant-Q 알고리즘의 페로몬 갱신식을 이용하였고 큰 페로몬 값을 가중치로 두어 안정된 매칭을 이룰 수 있었다. 매칭 M 의 함수 값 변수 S 는 수식(8)의 $sw + sm$ 을 $|sw - sm|$ 으로 바꾼다.

$$AQ = \frac{\prod_{i=1}^{n-1} (f(M) - i)}{(n-1)!} \cdot \frac{1}{(sm - sw)} \quad (9)$$

수식(9)에서 $|sm - sw| = 0$ 일 때,

$0 \leq |sm - sw| \leq n - 1$ 을 만족하며, AQ 는 수식(10)을 바탕으로 생성된다.

$$AQ = 2 \cdot \frac{\prod_{i=1}^{n-1} (f(M) - i)}{(n-1)!} \quad (10)$$

$f(M) < n$ 이면, $F = f(M)$ 이다.

($AQ=0$ 일 때, 모든 $f(M) < n$ 이다.)

$|sm - sw|$ 이 최소값을 가지고, $f(M) = n$ 이면 $F = n + |sm - sw|^{-1}$ (또는 $F = |n + 2$ if $sw = sm$)이다.

<표 1> 정의 10에 의거한 평가 결과

	sm	sw	sm+sw	sm-sw
M_1	16	55	71	39
M_2	58	15	73	43
M_3	21	37	58	16
M_{4-1}	42	43	85	1

4. 결 론

SMP에 Ant-Q 알고리즘을 새롭게 적용한 본 논문은 기존에 제시되었던 Gale-Shapley 알고리즘과 ACS알고리즘과의 성능의 비교 평가를 위해 다음과 같이 환경설정을 하였다.

$a = 0.1, b = 2, q_0 = 0.9, n = 20$ 으로 설정 하였다.

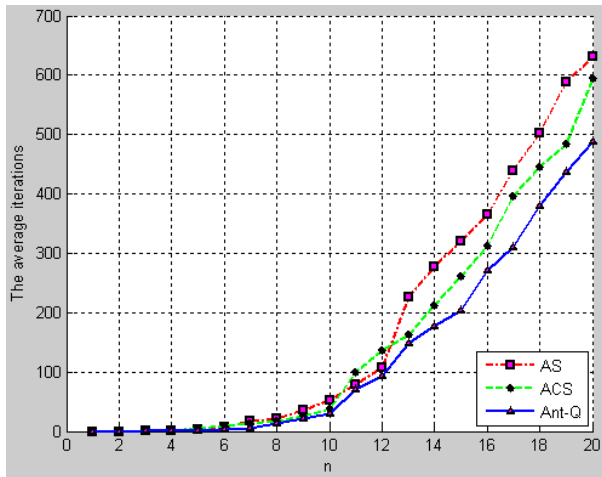
기존의 해법들과 비교를 통한 Ant-Q알고리즘의 성능 입증을 위해 m_i, w_j 의 각각의 수와 에이전트(Ant)의 숫자는 $n = 20, m = 15$ 로 하고 1000회 반복으로 실험을 실시하였고 SMP의 대표적 해법인 Gale-Shapley 알고리즘과 가장 최근에 연구되었던 ACS알고리즘[12]과 비교 평가 하였다. m_i 를 우선으로 하는 안정된 매칭과 w_j 를 우선으로 하는 안정된 매칭의 결과는 <표 2>와 같이 기존의 알고리즘들에 비해 성능 향상이 있음을 보여 주고 있다. <표 2>는 Ant-Q, Gale-Shapley[2], ACS(Ant Colony system)[12] 알고리즘을 적용한 매칭의 결과를 보여주고 있다. Ant-Q알고리즘은 본래 ACS보다 노드의 수가 많아질수록 성능값이 향상되기에[1] 본 논문에서도 기존의 ACS를 적용해 구한 SMP 해보다 성능의 향상이 있음을 알 수 있었다.

<표 2> 기존의 알고리즘과의 성능 평가(n=15,m=20)

	Ant-Q	Gale-Shapley	ACS
Men-optimum(sw)	67,53	68,69	69.12
Women-optimum(sm)	80.1	83.6	83.2
Egalitarian(sm+sw)	136	-	121
Sex-fair(sm-sw)	1	-	1

본 논문에서 제시한 Ant-Q 알고리즘은 Q-학습을 통해 SMP의 새로운 해법을 찾을 수 있었다.

(그림 3)은 Ant-Q 알고리즘의 매칭 진행 속도를 기존의 AS(Ant System)[15], ACS(Ant Colony System)[14]와 비교 평가한 결과를 보여주고 있다. Ant-Q가 반복 횟수가 늘어날수록 학습효과에 의해 기존의 AS, ACS 알고리즘에 비해 우수한 성능을 보여 주고 있다.



(그림 3) Ant-Q알고리즘의 성능 평가

본 논문에서는 SMP에 대한 연구에 대해 Ant-Q알고리즘을 적용하여 기존 Gale-Shapley알고리즘에서 찾았던 어느 한쪽만을 고려한 것과는 다르게 다양성에 초점을 두고 또 다른 2가지 형태의 매칭 방법을 제시하고 이에 따른 새로운 해법을 찾을 수 있었다. 그리고 이전에 제시된 ACS 적용 사례보다 Q-학습이라는 새로운 방법을 통해 새로운 SMP의 해를 찾는 것에 기여를 할 수 있었다.

하지만 기존과 마찬가지로 남자와 여자 수의 불균형 등과 같은 다양한 요소를 고려하지 않고 1:1이라는 제한적 환경에서 문제의 해법을 찾았기에 추후 연구에서는 m:n 형태의 각각의 비율이 다른 구성에서 안정된 매칭의 해를 찾을 수 있는 연구를 진행하고 본 논문에서 제시한 알고리즘의 성능을 평가해 보고자 한다.

참 고 문 헌

[1] Dorigo M. and L.M. Gambardella, "Ant-Q: A Reinforcement Learning Approach to Combinatorial Optimization" Tech. Rep. IRIDIA/95-01, Université Libre de Bruxelles, Belgium. 1995.
 [2] D.Gale and L.S.Shapley, "College admissions and the stability of marriage," American Mathematical Monthly, Vol.69, pp.9-15, 1962.

[3] D.Gusfield and R.W.Irving, The Stable Marriage Problem: Structure and Algorithms, The MIT Press, 1989.
 [4] An efficient algorithm for the "stable roommates" problem", Journal of Algorithms 6 (4): 577 - 595. Irving, Robert W, 1985.
 [5] Marilda Sotomayor, "The stability of the equilibrium outcomes in the admission games induced by stable matching rules", International Journal of Game Theory Vol.36, Numbers 3-4 621-640 0020-7276, Springer-Verlag, 2008
 [6] Robert W. Irving "Stable matching problems with exchange restrictions", Journal of combinatorial optimization Volume 16, Number 4 344-360 1382-6905, 2008
 [7] Bettina Klaus, Flip Klijn, Jordi Massó "Some things couples always wanted to know about stable matchings (but were afraid to ask)" Review of Economic Design Vol.11, Number 3 175-184 1434-4742, 2006
 [8] A. Colomi, M. Dorigo, and V. Maniezzo, "An Investigation of Some Properties of An Ant Algorithm." Proceedings of the Parallel Problem Solving from Nature Conference (PPSN 92), R. Määner and B. Manderick(Eds.), Elsevier Publishing, pp.509 - 520. 1992,
 [9] Caironi P.V.C. & M. Dorigo "Training and Delayed Reinforcements in Q-learning Agents" International Journal of Intelligent Systems, in press. (Also available as Tech. Rep. IRIDIA/94-14, Université Libre de Bruxelles, Belgium.), 1997.
 [10] Li Li, F. Qiao, Q. D. Wu, "ACO-based multi-objective scheduling of parallel batch processing machines with advanced process control constraints", The International Journal of Advanced Manufacturing Technology Vol.44, Numbers 9-10 985-994 1433-3015, Springer-Verlag, 2009.
 [11] B. Hölldobler and E.O. Wilson, "The ants". Springer-Verlag, Berlin, 1990.
 [12] 김현, 정태충, "ACS 알고리즘을 이용한 안정된 결혼 문제 해결에 관한 연구", 전자공학회 논문지 제 47권 CI편 제 6호 pp.68-74, 2010. 11.
 [13] M Dorigo, and L. Gambardella. Ant Colony System: A Cooperative Learning Approach To The Traveling Salesman Problem, IEEE transactions on Evolutionary Computing, 1(1):53-66, 1997.
 [14] M. Dorigo, V. Maniezzo, and A.Colomi. The Ant System: Optimization by A Colony of Cooperating Agents. IEEE Transactions on Systems, Man, and Cybernetics - Part B, Vol.26, No.2, pp.29 - 41, 1996.
 [15] R.W. Irving, D.F. Manlove and S. Scott, "The Hospitals/Residents Problem with Ties", in Proceedings of SWAT 2000, the 7th Scandinavian Workshop on Algorithm Theory, volume 1851 of Lecture Notes in Computer Science (Springer-Verlag, 2000), pp.259-271.



김 현

e-mail : kimhyun@khu.ac.kr
2005년 경희대학교 컴퓨터공학과(석사)
2011년 경희대학교 컴퓨터공학과 공학박사
관심분야: 멀티에이전트, 강화학습,
스마트 TV



정 태 충

e-mail : tchung@khu.ac.kr
1987년 KAIST 전산학과(공학박사)
1987년 KIST 시스템공학센터 선임연구원
1988년~현재 경희대학교 컴퓨터공학과
교수
관심분야: 인공지능, 기계학습, 메타탐색,
로봇축구