

중 분화 진화 알고리즘을 이용한 안정된 베이지안 네트워크 앙상블 구축

(Construction of Robust Bayesian Network Ensemble using
a Speciated Evolutionary Algorithm)

유 지 오 [†] 김 경 중 [†] 조 성 배 ^{**}
(Ji-Oh Yoo) (Kyung-Joong Kim) (Sung-Bae Cho)

요약 베이지안 네트워크는 불확실한 상황을 모델링하기 위한 확률 기반의 모델로서 확실한 수학적 토대를 가지고 있다. 베이지안 네트워크의 구조를 자동 학습하기 위한 연구가 많이 있었고, 최근에는 진화 알고리즘을 이용한 연구가 많이 진행되고 있다. 그러나 대부분은 마지막 세대의 가장 좋은 개체만을 이용하고 있다. 시스템이 요구하는 다양한 요구 조건을 하나의 적합도 평가 수식으로 나타내기 어렵기 때문에, 마지막 세대의 가장 좋은 개체는 종종 편향되거나 변화하는 환경에 덜 적응적일 수 있다. 본 논문에서는 적합도 공유 방법으로 다양한 베이지안 네트워크를 생성하고, 이를 베이스 규칙을 통해 결합하여 변화하는 환경에 적응적인 추론 모델을 구축할 수 있는 방법을 제안한다. 성능 평가를 위해 ASIA와 ALARM 네트워크에서 인공적으로 생성한 데이터를 이용한 구조 학습 및 추론 실험을 수행하였다. 다양한 조건에서 학습된 네트워크를 실험한 결과, 제안한 방법이 변화하는 환경에서 더욱 강건하고 적응적인 모델을 생성할 수 있음을 알 수 있었다.

키워드 : 베이지안 네트워크, 중 분화, 앙상블, 진화 연산, 강건한 추론

Abstract One commonly used approach to deal with uncertainty is Bayesian network which represents joint probability distributions of domain. There are some attempts to learn the structure of Bayesian networks automatically and recently many researchers design structures of Bayesian network using evolutionary algorithm. However, most of them use the only one fittest solution in the last generation. Because it is difficult to combine all the important factors into a single evaluation function, the best solution is often biased and less adaptive. In this paper, we present a method of generating diverse Bayesian network structures through fitness sharing and combining them by Bayesian method for adaptive inference. In order to evaluate performance, we conduct experiments on learning Bayesian networks with artificially generated data from ASIA and ALARM networks. According to the experiments with diverse conditions, the proposed method provides with better robustness and adaptation for handling uncertainty.

Key words : Bayesian Network, Speciation, Ensemble, Evolutionary Computation, Robust Inference

1. 서론

인공지능 분야에서 베이지안 네트워크(Bayesian network, BN)는 불확실성을 처리하기 위한 중요한 방법

중 하나로 부각되고 있다. 베이지안 네트워크는 현실 세계의 문제를 결합 확률 분포로 나타낸 모델로, 전문가의 지식을 쉽게 반영할 수 있는 장점이 있다. 이 모델은 방향성 비순환 그래프(directed acyclic graph, DAG)로 표현되는데, 이 그래프는 환경 변수를 나타내는 노드와 각 변수 간의 의존성을 나타내는 아크로 구성된다[1]. 적은 수의 노드가 주어지더라도 가능한 DAG 구조는 매우 많기 때문에, 주어진 데이터로부터 최적의 DAG 구조를 탐색하는 것은 매우 어렵다. 최근 진화 알고리즘을 이용하여 베이지안 네트워크를 학습시키는 연구가 많이 진행되고 있지만, 대부분 마지막 세대의 하나의 최

· 본 연구는 정보통신부 21세기 프론티어 연구개발사업인 '유비쿼터스 프론티어 사업단'의 지원으로 수행되었습니다.

[†] 학생회원 : 연세대학교 컴퓨터과학과
taiji391@sclab.yonsei.ac.kr
uribyul@sclab.yonsei.ac.kr

^{**} 종신회원 : 연세대학교 컴퓨터과학과 교수
sbcho@csai.yonsei.ac.kr

논문접수 : 2004년 7월 30일
심사완료 : 2004년 10월 26일

적해만 찾는 것이 보통이다[1-3].

환경이 변화했을 때 그에 맞추어 모델을 적응시키는 것은 비용이 많이 들거나 불가능한 경우도 있기 때문에, 최초로 강건한 모델을 학습하는 것이 좋다. 최적의 해 하나만을 찾는 일반적인 진화 알고리즘에서 강건한 해를 구하려면 실제 상황에서 있을지 모르는 노이즈까지 정확하게 반영할 수 있는 효과적인 평가 방법이 필요하다. 그러나 노이즈는 언제 나타날지 예측할 수 없기 때문에, 이러한 방법에는 많은 어려움이 있다. 만약 다른 특성을 가진 여러 개의 해를 동시에 사용할 수 있다면, 최적 해 하나만을 사용하는 것보다 더 강건한 성능을 기대할 수 있을 것이다. 특히, 베이지안 네트워크는 사람이 쉽게 해석할 수 있고 전문가의 지식을 쉽게 반영할 수 있기 때문에, 환경이 변화했을 때 적은 비용으로 구조를 적응적으로 변화시킬 수 있다.

현실 세계의 많은 문제들은 시간이 지나감에 따라 변화하고, 그 변화를 예측하기 어렵기 때문에 동적인 환경의 불확실성을 여러 휴리스틱한 방법을 통해 처리하려는 연구가 많다. 그런 방법 중의 하나로 진화 알고리즘이 있지만 일반적인 진화 알고리즘은 지역 해에 빠질 수 있기 때문에 변화하는 환경에 적응적인 해를 탐색하는데 필요한 다양성을 잃어버릴 우려가 있다. 이에 반해 종 분화 진화 알고리즘은 하나의 최적 해를 찾는 것이 아니라 다양성을 유지하면서 해를 탐색하기 때문에 다른 특성을 가진 여러 개의 최적에 가까운 해를 찾을 수 있다. 본 논문에서는 종 분화 진화 알고리즘의 하나인 적합도 공유 알고리즘을 사용하며, 각 개체간의 거리를 MDL(Minimum Description Length)[4]을 사용하여 측정한다. 측정된 거리는 적합도 공유 알고리즘에서 종을 구분하는 기준으로 사용된다. 다양한 해들을 결합하면 서로간의 상호 작용을 통해 하나의 최적 해를 사용하는 것보다 더 좋은 결과를 가져다 줄 수 있다.

본 논문에서는 적합도 공유를 이용하여 다양한 베이지안 네트워크를 생성하고, 이를 베이즈 규칙(Bayes rule)을 통해 결합하는 방법을 제안한다. 동적인 환경에서 적합도 공유를 통해 생성된 여러 베이지안 네트워크 중 일부는 특정 상황에서 잘못된 결과를 낼 가능성이 있지만, 상호 보완하여 올바른 결과로 수정될 수 있다. 제안한 방법의 유용성을 보여주기 위해서, 대표적인 벤치마크 문제인 ASIA 네트워크와 ALARM 네트워크에서 인공적으로 생성된 데이터를 사용하여 실험을 수행하였다.

본 논문의 구성은 다음과 같다. 2절에서는 제안한 방법을 이해하기 위한 기본적인 배경 지식을 다룬다. 3절에서는 종 분화가 어떻게 진화 알고리즘에 적용되고, 해들이 어떻게 결합되는지를 설명한다. 4절에서는 ASIA

와 ALARM 네트워크로부터 얻어진 데이터를 사용한 실험 과정과 분석 결과를 논의한다.

2. 배경

2.1 베이지안 네트워크

베이지안 확률 추론은 불확실한 상황을 표현하고 추론하는 모델로서, 베이지안 네트워크를 통한 방법이 대표적이다. 베이지안 네트워크에서 노드는 실제 환경 변수를 나타내고, 아크는 각 변수 간의 의존성을 나타낸다. 네트워크를 구축하기 위해서는 구조를 설계하고, 각 노드에 맞는 확률 분포를 정해주어야 하는데, 보통 구조는 전문가에 의해 설계되고, 확률 분포는 전문가가 계산하거나 혹은 주어진 데이터를 통해 계산된다. 네트워크를 학습한 후 어떤 상황에 대한 증거가 관찰되면 그 증거를 바탕으로 각 노드의 조건부 확률 테이블과 독립 조건을 이용, 베이지안 추론 알고리즘을 통해 각 노드의 상태에 대한 확률이 계산된다.

$$P\langle B, \theta_B \rangle = P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | Pa(x_i))$$

위 수식은 베이지안 네트워크를 나타낸 것이다. $\langle B, \theta_B \rangle$ 는 베이지안 네트워크의 구조 B 와 확률 변수 θ_B 를 나타내고 $P\langle B, \theta_B \rangle$ 는 네트워크의 모든 변수 간의 결합 확률 분포를 나타낸다. 베이지안 네트워크의 구조를 방향성 비순환 그래프 $B=(V, E)$ 라고 하면, $V=\{x_1, x_2, \dots, x_n\}$ 는 노드의 집합으로 실제 문제의 환경 변수를 의미하고, E 는 아크의 집합으로 각 변수 간의 의존 관계를 나타낸다. 각 변수 $x_i \in V$ 에 대해서, 조건부 확률분포는 $P(x_i | Pa(x_i))$ 로 나타낼 수 있고, 여기서 $Pa(x_i)$ 는 변수 x_i 의 부모 집합이다.

데이터로부터 베이지안 네트워크를 학습시키는 주요 접근 방법으로 여러 개의 네트워크 후보를 두고 score metric을 이용해 각 네트워크의 적합도를 평가하면서, 네트워크 후보를 탐색해나가는 방법이 있다[5]. 일반적으로 베이지안 네트워크를 학습하는 문제는 NP-hard 문제이기 때문에[6] 휴리스틱한 방법으로 학습시키는 연구가 있어왔다. 진화 알고리즘은 동적인 현실 세계를 최적화하는 문제에 대해 적합하기 때문에, 베이지안 네트워크를 진화로 학습시키려는 시도가 최근 진행되고 있다.

진화 연산을 적용시키기 위한 두 가지 방법으로 순서(Ordering) 가정을 두는 방법과 회복 연산자(Repair operator)를 이용하는 방법이 있다. 이 방법들은 네트워크를 진화시킬 때 DAG 구조에 위반되지 않는 네트워크를 생성하기 위해 제안되었다. 순서 가정이란 만약 노드 x_j 가 x_i 의 부모 노드라면, 입력 순서에서 x_j 는 반드시

시 x_i 보다 앞이어야 한다는 것이다. 이 경우, 베이지안 네트워크를 나타내는 각 개체의 길이는 노드의 수가 n 이라 했을 때 nC_2 이므로, 해공간은 2^{nC_2} 로 줄어들고 일반 진화 연산자를 적용하여 올바른 네트워크 구조를 생성할 수 있다. 그러나 노드 간 순서를 가정하지 않고 진화를 하게 되면, 일반 진화 연산자를 적용할 경우 잘못된 DAG 구조가 생성될 수 있다. 또한 해공간도 다음과 같이 커진다[7].

$$f(n) = \sum_{i=1}^n (-1)^{i+1} nC_i 2^{(n-i)} f(n-i), \quad f(0)=1, \quad f(1)=1$$

이 경우에는 노드의 순서를 가정하는 대신 회복 연산자를 사용하여 적합하지 않은 네트워크 구조를 적합한 구조로 바꿀 수 있다.

최근, 몇몇 연구자들은 불확실성을 다루기 위해 베이지안 네트워크의 효과적인 학습 및 모델링에 관한 연구를 진행하고 있다. Kim 등은 영상으로부터 3D 물체 정보를 계산하는 Expandable Bayesian Network(EBN)을 제안하였다[8]. 또한 Lam은 부분적으로 입력된 데이터를 이용하여 베이지안 네트워크를 환경에 맞게 개량하는 방법을 제안했다[4]. 진화 알고리즘을 이용해 베이지안 네트워크를 학습시키는 연구도 수행되었는데, Larranaga 등은 ASIA와 ALARM 네트워크를 대상으로 한 실험에서 진화 알고리즘의 각종 매개 변수(개체 수, 지역 최적화 방법, 교차 확률, 돌연변이 확률)를 변화시켜가며 성능 분석을 했다[1,9]. 또한 가장 최적의 노드 순서를 진화 알고리즘을 통해 탐색하는 연구도 수행하였다[2]. Wong 등은 MDL과 진화 프로그래밍(Evolutionary Programming, EP)을 기반으로 베이지안 네트워크를 학습시키는 새로운 방법을 제안하였다[3]. Li 등은 KL 거리(Kullback-Leibler distance)를 사용하여 베이지안 네트워크 구조와 데이터 사이의 거리를 계산하고, 회복 연산자를 사용한 진화 알고리즘을 통해 최적의 구조를 탐색하는 방법을 제안하였다[10]. 또한 Cotta 등은 여러 가지 하이브리드 교차 연산자를 사용하여 데이터로부터 베이지안 네트워크를 추출하는 방법을 제안하였다[11].

2.2 베이지안 네트워크 앙상블

여러 개의 모델을 결합하기 위해서는 각 모델의 특성이 다를수록 좋다. 그러나 일반적인 진화 알고리즘은 진화 과정 중 다양성을 잘 유지하지 못하기 때문에 다른 특성의 모델들을 생성하는데 어려움이 있다. 몇몇 연구자들은 진화를 진행하는 도중에도 다양성을 유지시키기 위해 진화 알고리즘 중 선택 과정을 변형하는 중 분화 알고리즘을 제안하였다. Anderson은 유전자형과 표현형을 공유하는 유전자 알고리즘을 통해 트랙을 따라 움직

이는 문제를 최적화할 수 있는 방법을 연구했다[12]. Cedeno 등은 Crowding과 유사한 알고리즘을 통해 다양한 해가 존재할 수 있는 해 공간에서 다중 해를 찾는 연구를 하였다[13]. Ghosh 등은 각 개체의 나이를 고려할 수 있도록 적합도 평가 함수를 변형하는 것을 제안하고, 이런 접근 방법이 나이를 고려하지 않는 진화 알고리즘보다 다양성을 유지할 수 있다는 점을 보였다[14].

앙상블은 각 개체간의 상호작용을 통하여 동적인 환경에 적응적인 모델을 만들 수 있다. Garg 등은 Schapire와 Freund의 Adaboost 알고리즘을 사용한 베이지안 네트워크 학습 방법을 발전시켰다[15]. 이 모델은 이산적, 연속적인 상태에서 정적, 동적인 베이지안 네트워크까지 고려한 모델이다. Larranaga 등은 세 가지 다른 종류의 베이지안 네트워크 구조 (Naïve Bayes, Interval Estimation Naïve Bayes (IENB), Joining Attributes in Naïve Bayes)를 결합하였다[16]. 이들의 연구에서 보여준 예측 정확도는 기존의 방법보다 평균 1.21%를 상회하는 결과를 보였다.

3. 중 분화된 진화 베이지안 네트워크 앙상블

그림 1은 진화 베이지안 네트워크 중분화의 전체적인 순서도이다. 이 시스템은 우선 각각의 BN을 임의의 구조로 초기화한다. 학습 데이터에 대한 베이지안 네트워크의 점수를 구하는 score metric 중 DPSM(General Dirichlet Prior Score Metric)이 다른 것에 비해 좋은 성능을 보이기 때문에[5], DPSM을 사용해 각 베이지안 네트워크의 적합도를 계산하였다. 적합도 공유에서 각 개체 간 거리는 계산하는데 비용이 적은 MDL를 사용

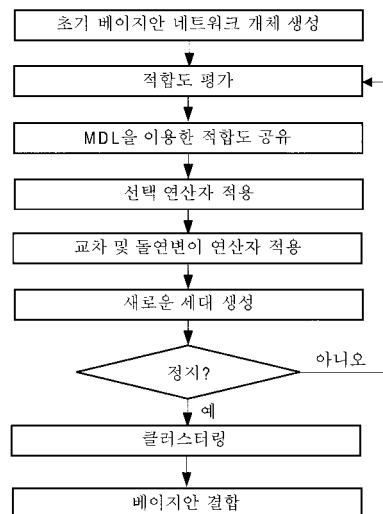


그림 1 중 분화 진화 베이지안 네트워크의 실행 순서도

하여 구현되었다. 적합도에 따라 80%의 비율로 선택 연산자를 통해 좋은 개체를 선택하고, 선택된 개체는 교차와 돌연변이 연산을 거치면서 새로운 세대를 형성한다. 미리 설정한 임계치보다 적합도가 높은 개체들이 일정 수가 탐색되거나, 최대 세대 수에 도달하면 진화 연산은 멈추게 되고, 클러스터링을 통해 마지막 세대에서 결합할 개체들을 선택한다.

3.1 염색체 표현

진화 알고리즘에서 각 개체의 표현 방법을 결정하는 것은 매우 중요한 문제이다. 베이지안 네트워크를 표현하는 방법은 여러 가지가 알려져 있는데 대표적으로 연결 행렬(Connection matrix)과 노드 순서(Variable order)가 있다. 연결 행렬로 표현할 때, 베이지안 네트워크는 다음과 같은 요소 c_{ij} 로 이루어진 $n \times n$ 연결 행렬 C 로 표현할 수 있다.

$$c_{ij} = \begin{cases} 1 & x_j \text{가 } x_i \text{의 부모 노드일 경우,} \\ 0 & \text{그렇지 않을 경우} \end{cases}$$

연결 행렬은 표현하기 쉽고 진화 연산을 쉽게 적용할 수 있기는 하지만, 추가적으로 회복(Repair) 연산이 필요하다. 한편 베이지안 네트워크 학습에서 노드의 순서가정을 통해 탐색 공간을 줄이는 것이 가능한데 환경 변수의 순서는 다음과 같이 정의할 수 있다. 베이지안 네트워크 구조는 노드 순서 L 로 표현할 수 있는데, 이때 길이는 n 이고, 요소 l_i 는 다음과 같은 규칙에 의해 표현된다.

$$l_j, l_i \in V, j < i \text{ if } l_j \in Pa(l_i)$$

순서를 가정한 방법은 탐색 공간을 줄여주지만, 다양한 구조를 생성하는데 제한이 있다.

본 논문에서는 위의 두 가지 표현 방식을 결합한 새로운 개체 표현 방법을 제안한다. 이 방법은 회복 연산자를 사용하지 않으면서, 다양한 베이지안 네트워크 구조를 탐색하는데 적절한 표현 방법이라고 할 수 있다. 제안한 표현 방법에서 베이지안 네트워크 구조는 변수의 순서를 나타내는 길이가 n 인 L 과 $n \times n$ 연결 행렬 C 로 표현된다. 각각의 요소 c_{ij} 는 다음과 같은 방법으로 표현되는데, 각 변수 사이에 아크가 존재하는지에 대한 여부를 표시한다.

$$c_{ij} = \begin{cases} 1 & i > j \text{이고 } x_j \text{와 } x_i \text{사이에 아크가 존재할 경우,} \\ 0 & \text{그렇지 않을 경우} \end{cases}$$

노드 순서 L 에 대한 정의는 기존의 방법과 동일하다. 그림 2는 베이지안 네트워크의 표현 예를 보인 것이다. 이 그림에서 노드의 순서는 입력 순서를 나타내고, 연결 행렬은 노드 간 연결 관계가 있는지 여부를 나타낸다. 예를 들어 c_{31} 이 '1'이고 x_1 가 x_3 보다 앞에 위치하고 있다면, x_1 에서 x_3 방향으로 아크가 결정된다.

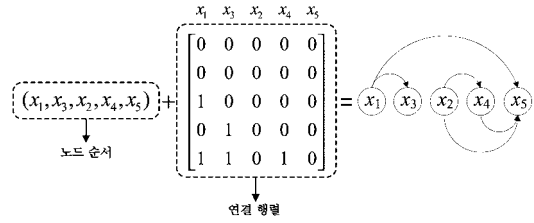


그림 2 베이지안 네트워크의 표현 예

3.2 유전 연산

각 개체의 적합도가 평가되면, 진화 알고리즘으로 일정 비율로 좋은 개체를 선별하고, 교차와 돌연변이 연산을 거친다. 본 논문에서는 순위 기반 선택(Rank-based selection) 연산자를 사용하였다. 이 연산자는 다음과 같은 정의에 의해 작동된다. I'_j 를 세대 t 에서 j 번째 개체라고 하고, $rank(g(I'_j))$ 를 적합도에 따른 각 개체의 순위, λ 를 개체의 수라고 할 때, 각 개체 I'_j 의 선택 확률 $p_{j,t}$ 는 다음과 같이 나타낼 수 있다.

$$p_{j,t} = \frac{\lambda - rank(g(I'_j)) + 1}{\lambda(\lambda + 1)/2}$$

교차 연산자는 개체 집단 안에 있는 두 베이지안 네트워크의 구조를 교환하는 역할을 한다. 최적의 변수 순서를 찾는 것은 NP-hard 문제이다. Larranaga 등은 TSP 문제를 대상으로 순서를 결정하는 여러 교차 연산자를 실험하였다[2]. 실험 결과에 따르면 순환 교차 연산자(Cycle crossover operator, CX)가 가장 좋은 성능을 보여주고 있었는데, 다른 교차 연산자는 개체 집단의 크기가 큰 경우에만 좋은 성능을 보였지만, 순환 교차 연산자는 적은 개체 수만으로도 좋은 성능을 유지했다. 순환 교차 연산자는 각 개체의 위치에 대응되는 상대 개체의 위치를 따라가며 사이클을 이루는 부분을 분리하고 이들을 교환하는 방식으로 교차를 수행한다. 예를 들어, 그림 3을 보면 각 개체로부터 2개의 사이클을 얻을 수 있고, 이 둘을 교환함으로써 교차 연산이 이루어지고 있음을 알 수 있다.

연결 행렬 중 왼쪽 아래의 삼각 부분만 보면 유용한 정보를 모두 얻을 수 있기 때문에 $n \times n$ 연결 행렬은 nC_2 의 길이로 된 문자열로 표현할 수 있다. 이렇게 표현된 연결 행렬은 1-점 교차 방식으로 교차 연산이 적용된다.

변수 순서의 돌연변이 연산은 전치 돌연변이 연산자(Displacement mutation operator, DM)가 사용되었다. Larranaga 등의 연구에서 순환 교차 연산자와 전치 돌연변이 연산자를 같이 사용했을 때 가장 성능이 좋았다 [2]. 전치 돌연변이 연산은 먼저 무작위로 부분 문자열

을 선택한 다음, 이 부분 문자열을 제거해서 다른 임의의 위치로 삽입하는 방식이다. 그림 4는 전치 돌연변이 연산의 예이다. 문자열 $(x_1, x_3, x_2, x_4, x_5)$ 가 있고 여기에서 무작위로 선택된 부분 문자열을 (x_4, x_5) 라고 하면, (x_4, x_5) 가 제거되어 (x_1, x_3, x_2) 를 얻는다. 그 후 임의로 선택된 위치가 x_3 이면 부분 문자열 (x_4, x_5) 가 그 위치로 삽입되고, 결국 문자열 $(x_1, x_3, x_4, x_5, x_2)$ 로 돌연변이 된다.

연결 행렬에 대한 돌연변이는 새로운 연결관계를 주거나 기존의 연결관계를 제거시킨다. 만일 연결 관계가 없다면 연결 행렬 요소의 값은 '0'일 것이고, 이 값을 '1'로 바꾸어줌으로써 새로운 연결관계가 생성된다. 반대로 연결관계가 이미 있다면, 연결관계를 제거한다.

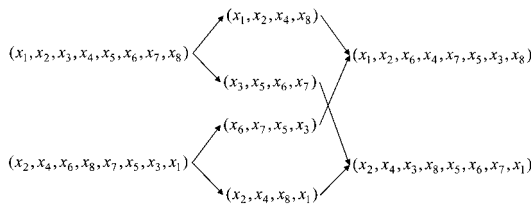


그림 3 순환 교차 연산의 예

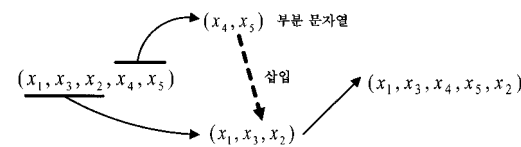


그림 4 전치 돌연변이 연산의 예

3.3 적합도 평가

데이터로부터 베이지안 네트워크를 학습하기 위해서, 많은 연구자들은 데이터에 대한 여러 가지 다른 가정을 바탕으로 다양한 종류의 score metric을 제안하였다. Score metric은 베이지안 네트워크가 주어진 데이터에 대해 얼마나 잘 설명하고 있는지를 평가한다. Yang 등은 UPSM(Uniform Prior Score Metric), CUPSM(Conditional Uniform Prior Score Metric), DPSM(Dirichlet Prior Score Metric), BDe(Likelihood-equivalence Bayesian Dirichlet Score Metric), MDL(Minimum Description Length) 등 다섯 가지 score metric의 성능을 비교한 바 있다[5]. 그리고 일반적으로 Dirichlet 분포 차수(Distribution order)가 10인 DPSM이 가장 좋은 결과를 낸다고 분석하였다. Dirichlet 분포를 가정으로 하는 DPSM은 다음과 같이 계산된다.

$$P(B, D) = P(B) \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(N_{ij}^*)}{\Gamma(N_{ij}^* + N_{ijk})} \times \prod_{k=1}^{r_i} \frac{\Gamma(N_{ijk}^* + N_{ijk})}{\Gamma(N_{ijk}^*)}$$

N_{ijk}^* 는 주어진 데이터베이스 D 에서 변수 x_i 의 상태가 k 이고($k=1, 2, \dots, r_i$), 그의 부모 $Pa(x_i)$ 의 상태는 j 인 경우($j=1, 2, \dots, q_i$)의 수이고, $N_{ij}^* = \sum_{k=1}^{r_i} N_{ijk}^*$ 이다.

N_{ijk}^* 는 Dirichlet 분포 차수(Distribution order)를 나타내고 사용자에게 의해 임의로 설정될 수 있다. 또한

$$N_{ij}^* = \sum_{k=1}^{r_i} N_{ijk}^*$$

Cooper 등은 어떻게 N_{ijk}^* 를 결정해야 하는지 언급하지 않았기 때문에[17], Yang의 연구 결과[5]에 따라, $N_{ijk}^* = 10$ 인 경우를 대상으로 실험하였다.

종 분화를 구현하는 방법은 여러 가지가 있는데, 여기서는 적합도 공유(Fitness sharing) 방식을 사용하였다. 적합도 공유는 개체들이 밀집된 지역에 있는 개체들의 적합도를 떨어뜨려서 다른 개체 간 적합도를 공유하는 방법이다. 적합도 공유는 진화 알고리즘이 해공간을 넓게 탐색할 수 있도록 하고, 결과적으로 좀더 다양한 종류의 개체를 생성할 수 있도록 해준다. 제한한 방법에서 각 개체는 베이지안 네트워크를 의미하는데 $\{B_1, B_2, \dots, B_{pop_size}\}$ 로 정의된다.

개체 B_i 의 적합도 f_i 가 주어지고, 공유 함수인 $sh(d_{ij})$ 가 있을 때, 공유 적합도 fs_i 는 다음과 같이 계산된다.

$$fs_i = \frac{f_i}{\sum_{j=1}^{pop_size} sh(d_{ij})}$$

공유 함수 $sh(d_{ij})$ 는 각 개체 B_i 와 B_j 간의 거리 d_{ij} 를 사용해 계산된다.

$$sh(d_{ij}) = \begin{cases} 1 - \frac{d_{ij}}{\sigma_s}, & 0 \leq d_{ij} < \sigma_s \\ 0, & d_{ij} \geq \sigma_s \end{cases}$$

여기서 σ_s 는 공유 반경이다. 공유 반경은 초기 개체 간 평균 거리를 이등분하여 사용하였다. 만약 각 개체 간의 거리가 σ_s 보다 크면 적합도를 공유하지 않고, 개체 간 거리가 σ_s 보다 작으면 적합도를 공유하게 된다. 자신과 적합도를 공유하는 개체가 많아질수록 공유 적합도는 자신의 원래 적합도보다 낮아지게 되고, 자신이 공유 반경 내에서 유일한 개체라면 공유 적합도는 원래 적합도와 같아지게 되므로, 각 공유 반경 내에 개체 수가 적을수록 다음 세대에 선택될 확률이 커지게 된다.

개체 B_i 와 B_j 사이의 거리에 대한 일반적인 거리 측정 함수는 없기 때문에, 구조적 차이를 거리로 사용할 수 있다. Lam은 기계 학습 분야에서 잘 정의되어 있는 MDL 이론을 이용하여 구조적 차이를 측정하는 척도를 정의했다[4]. B_i 와 B_j 에 대한 구조가 이미 정의되어 있

다면, B_i 와 B_j 간 구조의 차이를 측정할 수 있다. 만약 B_j 가 B_i 와 유사한 특성을 가지고 있다면, B_j 와 B_i 간 구조의 차이를 기술한 길이는 작을 것이고, 다른 특성을 가지고 있다면 구조의 차이를 기술한 길이는 클 것이다. 구조 차이를 계산하기 위해서 다음과 같은 기준을 따른다. 다음과 같은 기준에 의해 구조 차이를 계산하면, B_j 의 구조에 대한 정보와 B_j 와 B_i 간 구조의 차이를 기술한 정보를 통해 B_i 의 구조를 복구할 수 있다.

- 역방향 아크(B_j, B_i 모두 같은 위치에 아크가 있으나 뒤집혀있는 경우)
- B_i 에 빠진 아크(B_j 는 아크가 있으나 B_i 는 같은 위치에 아크가 없는 경우)
- B_i 에 추가된 아크(B_j 는 아크가 있으나 B_i 는 같은 위치에 아크가 없는 경우)

r, m, a 를 각각 B_j 의 역방향 아크 수, 빠진 아크 수, 추가된 아크 수라고 하면, 각 구조에는 $n(n-1)$ 개의 가능한 방향을 가진 아크가 존재할 수 있다. 이를 표현하기 위해서 $\log_2[n(n-1)]$ 비트가 필요하며, 이를 이용하여 거리 함수를 다음과 같이 나타낼 수 있다.

$$DL(B_i | B_j) = DL(B_j | B_i) = (r + m + a) \log_2[n(n-1)]$$

이 거리 함수는 계산이 복잡하지 않기 때문에, 적은 비용으로 계산될 수 있다.

3.4 다중 베이저안 네트워크의 결합

마지막 세대로부터 개체를 선별하기 위해서 단일 연결 클러스터링 기법을 사용한다. 결합에 사용될 개체의 수는 사전에 정의된 거리 임계치에 의해 자동으로 결정된다. 선별된 개체는 베이저안 결합 기법에 의해 결합된다. 베이저안 결합은 각 베이저안 네트워크의 오류 가능성을 염두에 두고, 오류에 대한 정보를 기록해두었다가 결합 결과에 반영하는 방법이다. 오류 가능성 테이블(Error Possibility Table, EPT)은 $M \times M$ 행렬로 표현될 수 있다(M 은 목적 변수의 상태 수이다). 이 때 행렬의 각 요소 $n_i^j(k)$ 는 현재 상태가 i 일 때, k 번째 베이저안 네트워크가 j 라는 상태를 출력하는 경우의 수이다. 이 오류는 $P(\text{target} \rightarrow \text{state } i | k\text{thBN, evidence } x)$ 에 의해 계산된다. K 가 결합에 사용될 베이저안 네트워크의 수라 하면, 신뢰도 함수 $BEL(i)$ 는 다음과 같이 나타낼 수 있다.

$$BEL(i) = \prod_{k=1}^K P(\text{target} \rightarrow \text{state } i | EPT_k, P(\text{target} \rightarrow \text{state } i | k\text{thBN, evidence } x))$$

결국 결합 함수 F 는 다음과 같이 정의된다.

$$F = \begin{cases} j & 0 < \alpha \leq 1 \text{ 일 때, } BEL(j) = \max_{i \in \{1, 2, \dots, M\}} BEL(i) \geq \alpha \text{ 인 경우} \\ \text{reject} & \text{그렇지 않을 경우} \end{cases}$$

여기서 α 는 임계치이다. 목적 노드의 상태 i 중에서 $BEL(i)$ 값이 α 보다 크면서 가장 큰 값인 상태를 결합

모델이 추론한 상태로 선택하게 된다. 임계치 α 가 커지면, 결합 모델에 대한 신뢰도는 향상될 수 있지만, 일반적으로 reject가 발생할 가능성이 높아진다.

4. 실험 결과

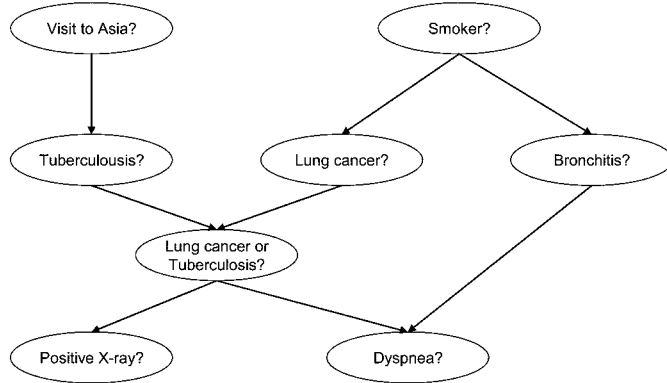
실험은 ASIA와 ALARM 네트워크에서 인공적으로 생성된 데이터를 사용하여 네트워크를 학습시킨 후 다양한 조건에서의 성능을 비교 평가하였다. ASIA 네트워크는 Lauritzen과 Spiegelhalter[18]에 의해 처음으로 소개되었다. ASIA 네트워크는 작은 규모의 가상적인 의학 지식을 가지고 설계되었는데, 환자의 여러 징후나 상태를 관찰하여 환자가 결핵에 감염되었는지, 폐암인지 혹은 기관지염인지에 대한 확률을 계산할 수 있다. ASIA 네트워크 구조는 환자가 아시아에 방문한 사실이 있을 경우 결핵에 걸릴 가능성이 높아지고, 호흡 곤란 증세를 보일 경우 결핵, 폐암, 기관지염일 가능성이 높아지는 등의 지식을 네트워크의 인과관계로 표현한 것이다. 그림 5(a)는 ASIA 네트워크의 구조를 나타낸 것이다.

ALARM 네트워크(그림 5(b))는 Beinlich 등이 만든 네트워크로, 병원의 중앙 조종실에서 환자에게 있을지 모를 마비 증상에 대한 기초적인 정보를 제공하는 네트워크이다[19]. 맥박, 호흡량 등 환자의 상태를 나타내는 수치 등을 통해 환자의 상태나 기계의 이상 유무, 오류 등의 가능성을 계산해주는 모델이라고 할 수 있다. ALARM 네트워크는 3가지 타입의 변수로 이루어져 있다. 진단 노드는 각각의 병이나 현상에 대한 유무를 상태 값으로 가지고 있어 네트워크가 추론할 수 있는 출력 부분이라고 할 수 있다. 측정 노드는 정량화된 입력을 받을 수 있는 노드로, 맥박, 호흡량과 같은 연속적으로 들어오는 입력 값들을 일정 범위로 구분하여 이산적인 값으로 바꾼 후 측정 노드의 상태에 입력한다. 중간 노드는 측정 노드로부터 입력된 데이터가 진단 노드의 출력으로 변해가는 과정에서 조절을 해주는 역할을 하는 노드이다.

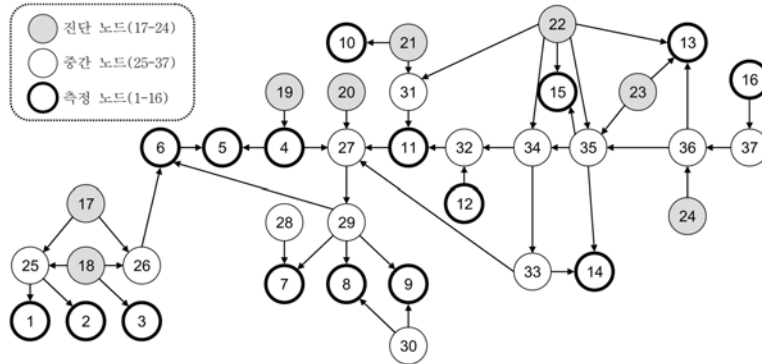
베이저안 네트워크로부터 데이터를 생성하기 위해 확률적 논리 추출(Probabilistic logic sampling) 기법을 사용하였다. ASIA의 경우 1000개의 케이스로 구성된 데이터를 생성하였고, ALARM은 3000개의 케이스로 구성된 데이터를 생성하였다. 진화 알고리즘의 개체 수는 50으로 설정했고, 최대 진화 세대 수는 ASIA의 경우 1000 세대, ALARM의 경우 5000 세대로 설정하였다. 선택 확률은 0.8, 교차 확률은 0.5, 돌연변이 확률은 0.01로 설정하였다.

4.1 ASIA 네트워크

평가를 위해 데이터에 포함된 1000개의 케이스 중 100개를 임의로 추출하여 900개의 케이스를 가진 학습



(a) ASIA 네트워크



(b) ALARM 네트워크

그림 5 벤치마크에 사용된 네트워크

데이터와 100개의 케이스를 가진 테스트 데이터로 나누어 학습을 수행하고 평가하였다. 네트워크의 규모가 작기 때문에 중 분화를 사용한 진화 알고리즘과 사용하지 않은 진화 알고리즘 모두 최적에 가까운 베이지안 네트워크를 찾을 수 있었다. 관찰되지 않고 다른 노드의 상태에 따라 상태가 추론이 되는 목표 노드는 “Tuberculosis,” “Lung Cancer,” “Bronchitis,” 그리고 “Lung Cancer or Tuberculosis”로 설정하였다. 이 4가지 노드는 환자가 어떤 병을 앓고 있을 확률을 나타낸다. 데이터 자체가 원본 ASIA 네트워크로부터 가상적으로 생성되었다고, 추론 결과는 위의 노드를 제외한 나머지 4개의 관찰된 노드의 상태만 보고 생성되기 때문에 원본 네트워크도 100% 정확한 추론을 못할 수 있다.

진화가 끝난 후, 마지막 세대의 베이지안 네트워크 개체들을 클러스터링한다. 임계치를 90으로 했을 때 중 분화를 통해 생성된 마지막 세대가 5개의 클러스터로 분리되었고, 각각의 클러스터로부터 개체 17, 49, 39, 29, 그리고 46을 선택하였다. 만약 클러스터 내에 속한 개체

가 2개 이상일 경우, 적합도가 가장 높은 개체가 그 클러스터의 대표 개체로 선택된다. 크기가 가장 큰 클러스터 내의 개체 중 49번 개체가 가장 좋은 적합도(-2123.92)를 나타내어 그 클러스터의 대표 개체로 선택되었다. 다른 대표 개체들은 -2445.52부터 -2259.57의 적합도 범위를 가진다. 이는 그림 6과 같이 선택된 대표 개체가 전체 개체가 -2827.20부터 -2123.92의 적합도 범위를 가지는데 비해 비교적 높은 적합도를 가지고 있으면서 각 클러스터로부터 선택된 개체이기 때문에 개체의 특성은 다르면서 높은 성능을 기대할 수 있다. 다른 4개의 클러스터에서 선택된 개체의 적합도가 마지막 세대에서 가장 좋은 개체의 적합도보다 낮아도 이 개체들이 결합하면 서로 상호 작용에 의해 좀더 강한 추론 결과를 기대할 수 있다.

원본 네트워크, 중 분화를 사용하지 않은 진화 알고리즘(그림 7(a)), 그리고 중 분화를 사용해 진화한 후 결합한 모델(그림 7(b)~(f))의 “Lung Cancer or Tuberculosis”에 대한 추론 적합도는 각각 98%, 98%, 그리고

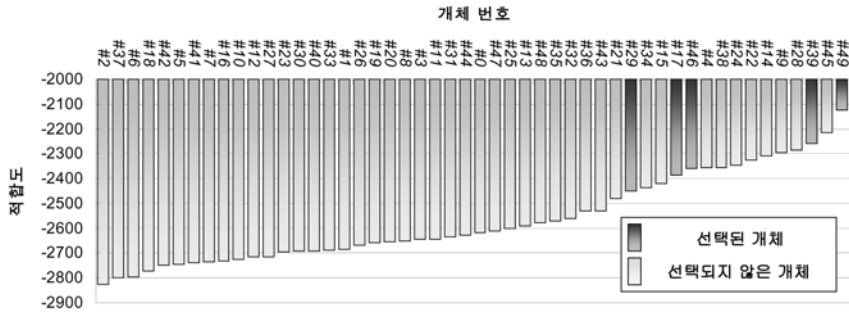
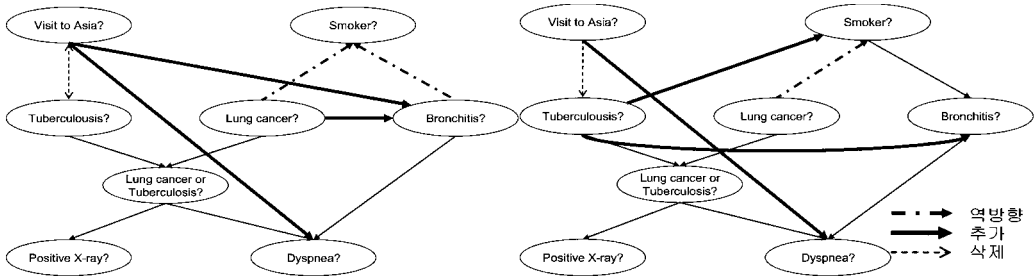
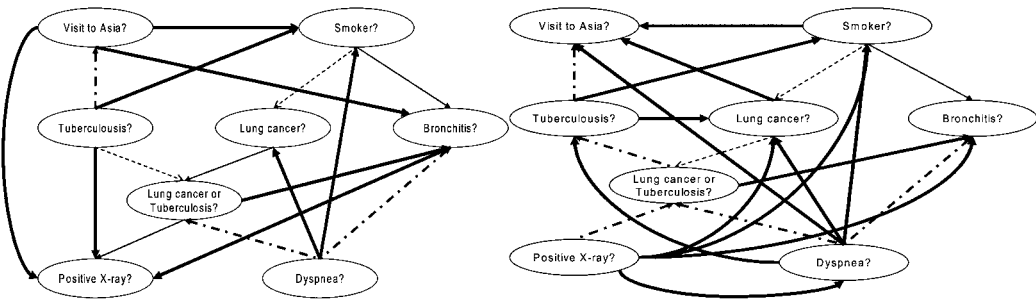


그림 6 각 개체들의 적합도 분포 (적합도가 높을수록 좋음)



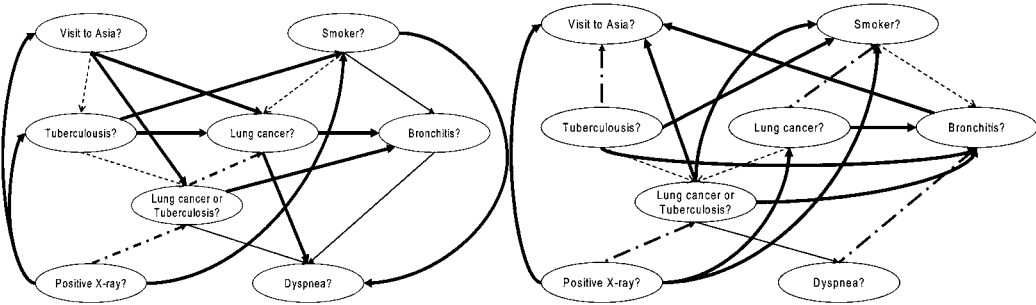
(a) 일반 진화 알고리즘 (적합도=-2128.26)

(b) 종 분화 49 번 개체(적합도=-2123.92)



(c) 종 분화 17 번 개체 (적합도=-2385.91)

(d) 종 분화 29 번 개체(적합도=-2445.52)



(e) 종 분화 39 번 개체 (적합도=-2259.57)

(f) 종 분화 46 번 개체(적합도=-2359.82)

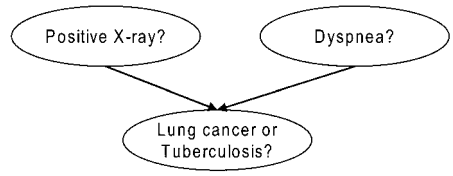
그림 7 일반 진화 알고리즘으로 진화된 개체 및 종 분화를 통해 진화된 5개의 개체 구조

표 1 “Visit to Asia”는 “No visit”, “Smoker”는 “Non-Smoker”, “Positive X-ray”는 “Abnormal” 그리고 “Dyspnea”는 “Present”인 경우, “Lung cancer or Tuberculosis” 노드의 결과, “Lung cancer or Tuberculosis”의 올바른 결과 값은 “True”이다.

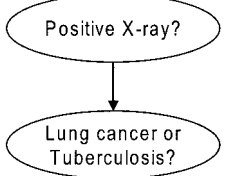
	원본 네트워크	일반 진화 알고리즘 최고 개체	중 분화 17번 개체	중 분화 29번 개체	중 분화 39번 개체	중 분화 46번 개체	중 분화 49번 개체	5개 BN의 베이지안 결합
	False	False	False	True	True	True	False	True
P(True)	0.494	0.471	0.118	0.709	0.676	0.576	0.464	0.503
P(False)	0.506	0.529	0.882	0.291	0.324	0.424	0.536	0.497

99%로 결합한 모델이 약간 더 좋은 결과를 보이고 있음을 알 수 있었다. 기타 다른 3개의 노드, “Tuberculosis”, “Lung Cancer”, “Bronchitis”에 대해서는 모든 모델이 같은 정확도를 나타내었다. 중 분화를 사용하지 않은 진화 알고리즘과 중 분화 알고리즘에서 가장 좋은 적합도를 가진 개체들은 원본 네트워크와 비슷한 구조를 가지고 있었다. 표 1은 결합한 모델이 더 좋은 결과를 내었던 경우에 대한 결과를 간략히 나타낸 것이다. 이 경우에서 원본 네트워크는 “True”일 확률이 0.494, “False”일 확률이 0.506으로 “Lung Cancer or Tuberculosis”가 “False”라고 실제 상황과 틀린 결과를 추론하고 있지만, 각 확률이 0.5와 근소한 차이를 보임으로써 어느 경우라고 단정지을 수 없는 불확실한 상황을 나타내고 있다. 그러나 일반 진화 알고리즘이나 중 분화 알고리즘의 최고 개체는 원본 네트워크와 같은 틀린 결과를 보여주고 있지만, 확률이 0.5와 어느 정도 차이가 있어 편향된 결과를 보여주고 있다. 이에 비해 결합한 모델은 올바른 결과를 추론할 뿐만 아니라 확률이 0.5와 근소한 차이를 보여주고 있어 다른 모델에 비해 원본 네트워크가 나타낸 불확실성을 가장 유사하게 모델링하고 있음을 알 수 있다.

표 1에서 나타났듯이, 원본 네트워크, 진화 알고리즘에서 나온 최고 적합도 개체들이 틀린 결과를 내는 경우에도, 29번, 39번, 46번 개체들은 정확한 결과를 추론하고 있다. 또한 29번 개체에서는 정확한 결과인 true가 될 확률이 가장 높게 나타나고 있다(0.709). 틀린 결과를 추론한 다른 구조들에 비해서 이들 3개 개체의 구조를 살펴보면 공통적인 부분이 있음을 알 수 있는데, “Positive X-ray?” 또는 “Dyspnea?” 노드가 “Lung Cancer or Tuberculosis”의 부모 노드라는 점이다. 그림 8은 이 네트워크들의 공통 구조를 간략히 나타낸 것이다. 이 부분은 원본 네트워크에 나타난 구조와 다르기 때문에 적합도가 낮게 평가되지만, 데이터에 대해 최적화된 적합도를 가진 구조가 가진 일부 경우에 대한 오류를 정정하고 있다고 볼 수 있다. 즉 중 분화를 통해 나온 대안적인 해들로부터 최적화된 구조가 편향된 결과를 추론하는 부분을 보완한다는 것이다. 실험의 경우



(a) 중 분화 29번 개체



(b) 중 분화 39번과 46번 개체

그림 8 정확한 예측을 한 개체들의 공통 구조

결합된 네트워크 중 가장 최적화된 네트워크에서 발생하는 오류를 다른 대안적인 3개의 네트워크가 상호작용을 통해 바로 잡은 경우라고 볼 수 있다.

정확도 분석은 베이지안 네트워크, 분류 트리, 회귀 트리(Regression Tree) 등의 모델을 평가할 때 쓰는 가장 일반적인 방법이다. 그러나 정확도만을 분석하는 방법의 최대 문제점은 각 모델이 예측하는 신뢰 정도를 측정할 수 없다는 점이다. 예를 들어 목적 상태에 대한 예측 확률이 0.51인 경우와 1.0인 경우, 정확도 분석에서는 모두 정확히 예측한 것으로 나타나지만, 신뢰 수준을 분석하면 0.51인 경우보다 1.0인 경우가 더 좋은 성능을 낸다고 볼 수 있다. Good은 모델의 신뢰 정도를 평가할 수 있는 정보 보상(Information reward) 척도를 개발했는데[20], 이에 대한 정의는 다음과 같다.

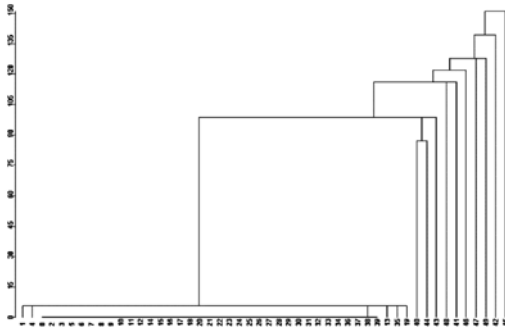
$$IR = \sum_i [1 + \log_2 P(x_i = v)]$$

여기에서 i 는 시험 데이터에 있는 케이스 번호를 의미하고, v 는 i 번째 케이스에 대한 실제 값, $P(x_i)$ 는 학습모델로부터 얻어진 실제 값에 대한 확률을 의미한다.

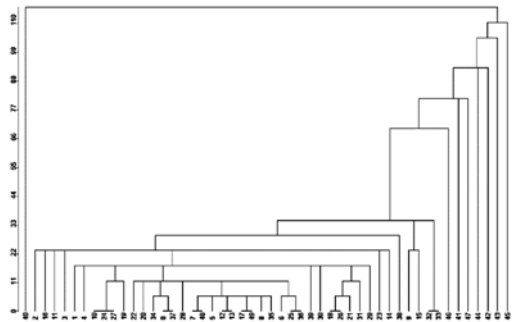
중 분화 베이지안 네트워크를 결합한 모델은 일부 노드의 상태가 관찰되지 않은 상황에서도 예측 가능성이

표 2 “Dyspnea”가 관찰되지 않았을 경우, “Bronchitis”의 정보 보상 정도. 총 10번의 실험 값의 평균이다.

원본 네트워크	일반 진화 알고리즘의 최고 적합도 개체	중 분화 알고리즘의 최고 적합도 개체	중 분화 베이지 안네트워크의 결합
-2.08	-3.14±0.18	-3.02±0.22	0.86±0.24



(a) 일반 진화 알고리즘



(b) 중 분화 진화 알고리즘

그림 9 덴드로그램 비교 (ASIA 네트워크) - x축은 개체 번호, y축은 개체간 거리

좋게 나타났다. 표 2는 결합한 모델이 가장 좋은 성능을 보이는 상황에서의 정보 보상 정도를 나타낸 것이다. 만약 보상 값이 높게 나타나면 추론이 잘되는 것을 의미한다. 중 분화를 통해 나온 가장 좋은 개체의 보상 값은 일반 진화 알고리즘을 통해 나온 가장 좋은 개체의 보상 값과 비슷하지만, 중 분화를 통해 나온 개체들을 결합한 모델은 정보 보상 정도가 원본 네트워크와 다른 모델보다 더 좋게 나타나고 있음을 볼 수 있다.

덴드로그램은 각 개체간의 연결 관계를 트리 형식으로 나타낸 그래프로 y 축은 유사한 정도, 즉 각 개체 간 거리를 의미하고, x 축은 각 개체 번호를 나타낸다. 그림 9는 각 알고리즘의 마지막 세대의 개체들을 덴드로그램으로 나타낸 것이다. 그림 9(a)에서 많은 개체들이 똑 같은 개체로 수렴해 하나의 클러스터를 형성하고 있음을 알 수 있다. 그러나 그림 9(b)에 있는 개체들은 비교적 여러 개의 클러스터를 형성하며 진화했음을 알 수 있어 중 분화 알고리즘이 다양한 특성을 가진 해를 탐색할 수 있음을 알 수 있다.

4.2 ALARM 네트워크

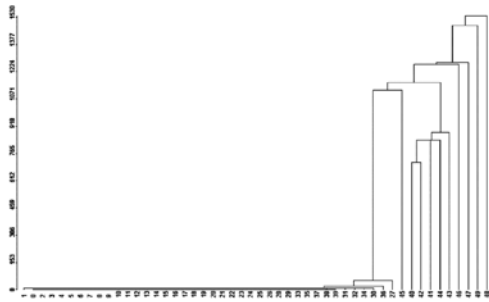
ALARM 네트워크는 노드의 개수가 37개로 노드가 8개인 ASIA 네트워크보다 규모가 훨씬 크고, 그만큼 학습 알고리즘이 탐색해야 하는 해공간도 커지게 된다. 37개의 노드 중, 10개의 노드는 관찰되지 않은 목표 노드로 두고, 나머지 노드는 관찰할 수 있는 노드로 설정하였다. 평가를 위해 데이터에 포함된 3000개의 케이스 중 600개를 임의로 추출하여 2400개의 케이스를 가진 학습 데이터와 600개의 케이스를 가진 테스트 데이터로 나누어 학습을 수행하고 평가하였다. 실험 결과는 총 5회의

임의 추출한 데이터를 사용하여 각각 2회씩 실험을 수행한 결과를 평균한 것이다. 표 3은 분류 정확도를 비교한 것이다. 일반 진화 알고리즘으로 생성된 개체가 중 분화 알고리즘의 개체보다 더 높은 정확도를 나타내고 있지만, 중 분화 알고리즘을 통해 생성된 개체를 결합하면, 오히려 높은 정확도를 나타내고 있음을 알 수 있다. 또한, 결합한 모델이 표준편차도 제일 작게 나타나 성능이 가장 안정적임을 알 수 있다. 그림 10은 일반 진화 알고리즘과 중 분화 알고리즘의 차이를 보여주고 있다. ASIA 네트워크와 마찬가지로 중 분화 알고리즘에서 생성된 개체가 비교적 여러 개의 클러스터를 형성하고 있어 더 다양한 해를 탐색하고 있음을 알 수 있다.

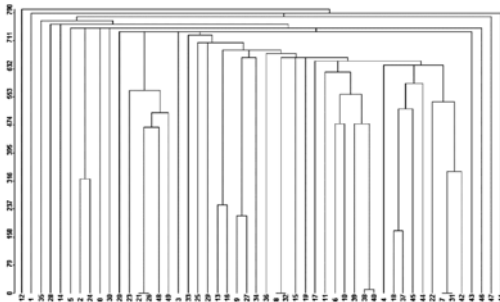
그림 11은 관찰 가능한 노드의 수가 변할 때 정보 보상 정도가 어떻게 변화하는지 나타낸 그래프이다. 관찰 가능한 노드가 적어질수록 불확실성은 증가하고, 이 때 정보 보상 정도의 변화가 적은 것이 불확실한 상황에 적응적이라고 할 수 있다. 관찰 가능한 측정 노드 11개(노드 번호 1, 2, 3, 5, 7, 8, 9, 10, 11, 14, 15)를 선정하고 관찰 가능한 노드의 수를 변화시켜가며 정보 보상 정도를 평가하였다. 원본 네트워크와 일반 진화 알고리즘은 관찰 가능한 노드가 적어지면 정보 보상 정도가 떨어지는데 반해, 제안한 방법은 일정한 성능이 계속 유지되면서 불확실한 상황에서 오히려 다른 모델보다 나은 성능을 보이고 있음을 알 수 있다. 이를 통해 결합한

표 3 “Hypovolemia” 노드에 대한 추론 정확도

일반 진화 알고리즘	중 분화 알고리즘	중 분화 개체의 결합
96.78±1.9	96.7±1.16	97.33±0.097



(a) 일반 진화 알고리즘



(b) 중 분화 진화 알고리즘

그림 10 덴드로그램 비교(ALARM 네트워크) - x축은 개체 번호, y축은 개체간 거리

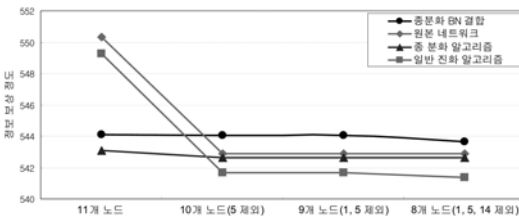


그림 11 관찰 가능한 노드 수에 따른 “Anaphylaxis” 노드에 대한 정보 보상 정도(ALARM 네트워크)

개체들이 상호 작용을 통해 불확실한 상황에서도 적응적인 추론을 할 수 있음을 알 수 있다.

5. 결론

본 논문에서는 불확실한 환경에 적응적인 베이지안 네트워크를 학습하기 위해 중 분화를 통해 얻어진 베이지안 네트워크들을 베이지안 결합 방법을 사용하여 결합하는 모델을 제안하였다. ASIA와 ALARM 네트워크에서 생성한 데이터를 이용한 실험 결과에서 제안된 방법이 오직 하나의 해만을 찾는 일반 진화 알고리즘보다 좋은 결과를 내고 있음을 확인할 수 있었다. 중 분화를 통해 생성된 네트워크를 결합하는 것이 각자를 따로 합

용하는 것보다 더 좋은 결과를 기대할 수 있다는 점도 확인했다. 즉, 결합에 사용된 베이지안 네트워크가 정확하지 않은 성능을 보이는 경우에도, 결합한 네트워크는 각 베이지안 네트워크가 출력하는 결과들의 오류를 고려하여 보다 정확한 추론을 할 수 있다는 것이다. 또 제안한 방법이 관찰 가능한 정보가 부족하여 불확실한 정보가 입력되더라도 적응적으로 추론할 수 있음을 알 수 있었다.

실제 세계는 동적으로 변할 뿐만 아니라 그 변화를 예측하기 어렵기 때문에 많은 불확실성이 내포되어 있다. 그렇기 때문에 환경이 동적으로 변할 때 그에 맞추어 적응적인 추론을 할 수 있는 모델을 학습하는 것이 중요하고, 본 논문에서 제안한 방법은 기존의 방법에 비해 적응적인 학습이 가능함을 알 수 있었다. 향후 연구로는 제안된 방법을 실제 환경에서 일어날 수 있는 문제에 적용시켜 보는 것이다. 특히 로봇 네비게이션이나 상황 인식, 사용자 행동 모델링 등의 문제에 적용시킬 계획이다.

참고 문헌

- [1] P. Larranaga, M. Poza, Y. Yurramendi, R. H. Murga, and C.M.H. Kuijpers, "Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 912-926, 1996.
- [2] P. Larranaga, C.M.H. Kuijpers, R. H. Murga, and Y. Yurramendi, "Learning Bayesian network structures by searching for the best ordering with genetic algorithm," *IEEE Trans. on Systems, Man and Cybernetics-Part A*, vol. 26, no. 4, pp. 487-493, 1996.
- [3] M. L. Wong, W. Lam, and K. S. Leung, "Using evolutionary programming and minimum description length principle for data mining of Bayesian networks," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 2, pp. 174-178, 1999.
- [4] W. Lam, "Bayesian network refinement via machine learning approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 240-251, 1998.
- [5] S. Yang, and K.-C. Chang, "Comparison of score metrics for Bayesian network learning," *IEEE Trans. on Systems, Man and Cybernetics-Part A*, vol. 32, no. 3, pp. 419-428, 2002.
- [6] D. M. Chickering, D. Geiger, and D. Heckerman, "Learning Bayesian networks is NP-complete," *Learning from Data: Artificial Intelligence and Statistics V (D. Fisher, H. Lenz (Eds.))*, Springer-Verlag, pp. 121-130, 1996.

- [7] R. W. Robinson, "Counting unlabeled acyclic digraphs," *Lecture Notes in Mathematics 622: Combinatorial Mathematics V*, pp. 28-43, 1977.
- [8] Z. Kim, and R. Nevatia, "Expandable Bayesian networks for 3D object description from multiple views and multiple mode inputs," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 769-774, 2003.
- [9] R. Etxeberria, P. Larranaga and J. M. Picaza, "Analysis of the behaviour of genetic algorithms when learning Bayesian network structure from data," *Pattern Recognition Letters*, vol. 18, no. 11-13, pp. 1269-1273, Nov 1997.
- [10] G. Li, F. Tong, and H. Dai, "Evolutionary structure learning algorithm for Bayesian network and penalized mutual information metric," *Proc. of IEEE Int. Conf. on Data Mining*, pp. 615-616, 2001.
- [11] C. Cotta, and J. Muruzabal, "Towards a more efficient evolutionary induction of Bayesian networks," *Lecture Notes in Computer Science 2439*, pp. 730-739, 2002.
- [12] H. C. Andersen, *An Investigation into Genetic Algorithms, and the Relationship Between Speciation and the Tracking of Optima in Dynamic Functions*, Honours thesis, Queensland University of Technology, Brisbane, Australia, November 1991.
- [13] W. Cedeno, and V. R. Vemuri, "On the use of niching for dynamic landscapes," *IEEE Int. Conf. on Evolutionary Computation*, pp. 361-366, 1997.
- [14] A. Ghosh, S. Tstutsui, and H. Tanaka, "Function optimization in non-stationary environment using steady state genetic algorithms with aging of individuals," *IEEE Int. Conf. on Evolutionary Computation*, pp. 666-671, 1998.
- [15] A. Garg, V. Pavlovic and J. M. Rehg, "Boosted learning in dynamic Bayesian networks for multimodal speaker detection," *Proc. of the IEEE*, vol. 91, no. 9, pp. 1355-1369, 2003.
- [16] V. Robles, P. Larranaga, J. M. Pena, E. Mena-salvas, M. S. Perez, and V. Herves, "Bayesian networks as consensed voting system in the construction of a multi-classifier for protein secondary structure prediction," *Artificial Intelligence in Medicine*, 2004. (in press)
- [17] G. F. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data," *Machine Learning*, vol. 7, pp. 299-347, 1992.
- [18] S. L. Lauritzen and D. J. Spiegelhalter, "Local computations with probabilities on graphical structures and their applications on expert systems," *Journal Royal Statistical Society B*, vol. 50, no. 2, pp. 157-224, 1988.
- [19] I. A. Beinlinch, H. J. Suermondt, R. M. Chavez and G. F. Cooper, "The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks," *Proc. of the Second European Conf. on Artificial Intelligence in Medicine*, pp. 247-256, 1989.
- [20] I. Good, "Rational decisions," *Journal of the Royal Statistical Society B*, vol. 14, pp. 107-114, 1952.



유 지 오

2004년 2월 연세대학교 컴퓨터과학과 졸업(학사). 2004년 3월~연세대학교 컴퓨터과학과 석사과정 재학중. 관심분야는 인공지능, 진화연산, 베이저안 네트워크, 컨텍스트 인식

김 경 중

정보과학회논문지 : 소프트웨어 및 응용 제 31 권 제 1 호 참조

조 성 배

정보과학회논문지 : 소프트웨어 및 응용 제 31 권 제 1 호 참조