

# 베이지안 망

김병희 (서울대학교 컴퓨터공학부)

## 분석자 서문

본 고에서는 확률적 모델링 및 추론 도구로서 과학기술 분야에서 널리 사용되고 있는 베이지안 망에 대해 살펴본다. 범용 도구로서 주목 받고 있는 베이지안 망의 기본 개념 및 특성과 함께 베이지안 망을 이용한 추론 알고리즘과 다양한 응용 사례를 정리하며, 관련된 다른 모델링 기법과의 연관성을 살펴본다. 베이지안 망의 개념을 처음 접하는 비전문가에서부터 활용을 고려하는 연구자에 이르기까지 다양한 배경을 가진 독자에게 적절한 깊이와 여러 예제를 통해 알기 쉽게 베이지안 망을 소개하는 분석 자료가 될 것이다.

## 1. 개요

베이지안 망(Bayesian network)은 지난 20~30 여 년 동안 다양한 분야의 과학자와 공학자의 주목을 받아왔다. 컴퓨터공학 분야에서는 인공지능적 측면에서, 통계학에서는 확률 그래프 모델 중 하나로써 폭넓은 연구 결과와 응용 사례가 축적되어 왔다. 집단유전학, 정보이론, 음성 인식, 컴퓨터 비전, 웹, 의학적 진단 등의 분야에서는 베이지안 망의 제한적 모델로 해석할 수 있는 다양한 방법론이 개발되어 왔다.

그렇다면 베이지안 망이란 무엇이고, 여러 분야에 폭넓게 적용되는 이유는 무엇일까? 베이지안 망은 주어진 상황에 대한 체계적인 확률 정보를 조직적이고 부분적으로 다룰 수 있는 기법이 일관되게 정리된 방법론이며, 구축된 모델을 활용하여 추론하기 위한 다양한 알고리즘이 갖추어져 있다. 이러한 점에서 베이지안 망은 불확실성 하에서의 모델 구축 및 추론이 필요한 문제라면 항상 적용할 수 있는 범용 도구로서 자리잡아 왔다. 수많은 응용문제가 베이지안 망에서의 추론 문제로 귀결되며 새로운 알고리즘을 고안할 필요 없이 베이지안 망 알고리즘을 활용하여 해결할 수 있다.

기술적으로 설명하자면, 베이지안 망은 기존의 확률 및 통계 기법으로는 다루기 힘든 고차원의 확률 분포를 간결하게 표현할 수 있다. 그림 1 에는 6 개의 이진 변수로 구성된 소규모 망이 표현되어 있다. 모든 베이지안

망은 방향성-비순환-그래프(구조라고 칭함)와 조건부 확률표(CPT, conditional probability table)의 집합으로 구성된다. 구조 상에서 노드는 변수에 해당하고 연결선은 확률적 독립성에 대한 표현으로 해석된다. 연결선은 직접적인 인과적 영향을 표현하는 것으로도 해석할 수 있다. 베이지안 망의 각 변수에는 각각의 CPT 가 부여되며, 변수와 그 부모 노드 간의 관계를 정량적으로 표현한다.

베이지안 망의 주된 특징으로 하나의 베이지안 망에 의해 부여되는 제한조건을 만족하는 확률분포가 유일하다는 점에서 일관성과 완벽성을 꼽을 수 있다. 그림 1 의 베이지안 망은 6 개 변수로 표현할 수 있는 64 가지 조합 전체에 대해 단일한 확률 분포를 결정한다. 이 확률 분포만으로 망의 변수에 의해 표현 가능한 모든 사건에 대해 확률을 계산할 수 있다.

베이지안 망의 또 다른 특징은 전체 확률 분포를 명시적으로 계산할 필요 없이 다양한 확률을 계산할 수 있는 효율적인 알고리즘이 존재한다는 점이다. 특히, 망의 형태에 관계 없이 이러한 알고리즘을 적용할 수 있다. 다만, 알고리즘의 효율성 및 정확도는 망의

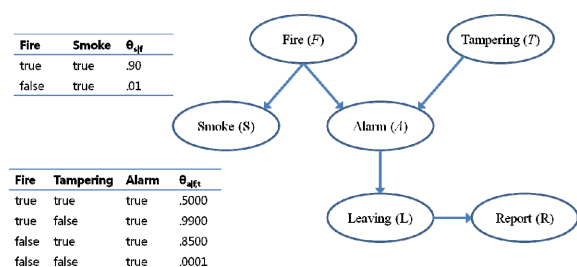


그림 1. 베이지안 망 및 조건부 확률표의 예

형태 및 특정 질의에 영향을 받는다. 흥미롭게도 유전학, 신뢰성 분석, 정보 이론 등과 같은 분야에서 과학자들이 개발해온 알고리즘들이 넓게 보면 모두 베이지안 망에 포함된다. 본 분석자료를 통해 이러한 연관성을 관련 연구자들에게 인식시키고, 각 분야의 문제 해결에 지난 수십 년 간 개발된 다양한 추론 기법을 적용할 수 있도록 베이지안 망을 소개하고자 한다[1].

## 2. 인과성 및 독립성

베이지안 망이 대규모 확률분포를 간결하게 표현하는 방식에 대해 살펴보자. 그림 1에서 표현한 예와 같이 조건부 확률표(CPT)는 전체 확률 분포에서 변수 간의 확률 분포에 대한 제한조건을 제공한다. 예를 들면,  $Pr(S=true|F=false)=0.01$  이다. 그러나, 이 조건만으로는 단일한 확률 분포가 정의되지 않으며, 베이지안 망의 그래프 구조까지 고려해야 한다.

베이지안 망을 구성하는 모든 변수는 부모 변수가 정해진 경우 자신의 자손이 아닌 변수와는 독립이라고 가정한다(조건부 독립). 그림 1에서 살펴보면, 변수  $L$ 은 그 부모  $A$ 가 알려진 경우 자손이 아닌  $T, F, S$ 와는 독립이다. 이러한 조건부 독립 조건은 베이지안 망에서의 마코프 가정(Markov assumption)으로 알려져 있으며, 조건부 확률표로 표현되는 수리적 제한조건이 결합되어 단일한 확률 분포를 정의하게 된다.

베이지안 망에서  $X \rightarrow Y$ 와 같이 연결되는 경우  $X$ 는  $Y$ 의 직접적인 요인으로 해석할 수 있으며 마코프 가정은 ‘변수의 직접적 요인이 알려진 경우 각 변수는 결과 이외의 변수에 독립’이라고 재해석된다.

베이지안 망으로 표현되는 확률 분포에서는 마코프 조건에 따른 독립성 이외에도 변수간 독립성을 추가로 확인할 수 있다. 베이지안 망의 그래프 구조상에  $d$ -분리( $d$ -separation)<sup>1</sup>라는 테스트를 수행하는 것만으로 쉽게 변수들 간의 조건부 독립성을 확인할 수 있다.  $D$ -분리의 기본 개념은 두 변수  $X, Y$ 를 연결하는 모든 경로가 변수  $Z$ 에 의해 막혀

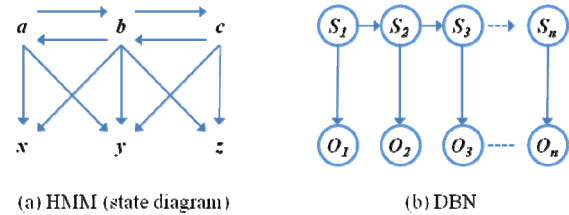


그림 2. 은닉변수모델(HMM)의 예 및 동등한 동적 베이지안 망(DBN)

있는 경우,  $X$ 와  $Y$ 는  $Z$ 가 주어졌을 한 조건부 독립이라는 것이다. 그림 1에서 경보기가 울린 상황( $A$  값이 알려진 경우)에서  $S$ 와  $T$ 가 연관이 있음을 다음의 경로를 살펴보면 확인할 수 있다 -  $\alpha: S \leftarrow F \rightarrow A \leftarrow T$ . 우선 화재는 연기의 직접적 요인이므로( $S \leftarrow F$ ), 연기가 관측된 경우 화재가 일어났을 가능성이 커진다. 화재 가능성이 커지면 경보기가 울린 요인으로서 화재가 더욱 부각되고 이에 따라 무단침입의 가능성은 낮아진다( $F \rightarrow A \leftarrow T$ ). 따라서,  $S$ 와  $T$ 를 잇는 경로는 열려 있으며 두 변수는 상호 연관성이 있음을 확인할 수 있다.

$D$ -분리를 이용한 독립성 테스트 시, 연결선에 대한 해석은 반드시 인과성일 필요는 없으며, 마코프 가정만으로도 두 변수 간에 가능한 모든 경로가 막혀 있는 경우 두 변수 간의 독립성을 확인할 수 있다.

$D$ -분리 테스트를 이용하면 다양한 분야에서 특화되어 개발된 확률 모델에 대해 이미 증명된 결과를 유도하는 것이 가능하다. 대표적인 예로 은닉 마코프 모델(HMM, hidden Markov model)을 살펴보자. HMM은 상태는 관측되지 않고 출력만이 관측되는 동적 시스템 모델에 널리 사용되는데, 시간적 패턴 인식이 필요한 음성인식, 필기인식, 동작 인식 등의 분야에서 활용되었으며, 생물정보학의 다양한 문제 해결에도 적용되었다. 그림 2a에 예로 든 HMM에는 세 개의 상태 ( $a, b, c$ )와 세 개의 출력( $x, y, z$ )으로 표현되는 시스템의 모델이 표현되어 있다. 각 상태간 전환 및 특정 상태에서 값이 출력될 확률이 그래프 구조에 추가되어 HMM 모델을 구성한다. 이 HMM은 그림 2b와 같은 베이지안 망으로 표현할 수 있다. 각 변수  $S_i$ 는 시간  $t$ 에서  $a, b, c$ 의 세 상태를 가질 수 있으며, 변수  $O_i$ 는 시간  $t$ 에서 출력될 수 있는 세 값  $x, y, z$ 를 표현한다. 이 망에  $d$ -분리 테스트를 적용하면 다음과 같은 HMM

<sup>1</sup>  $d$ 는 directed를 의미

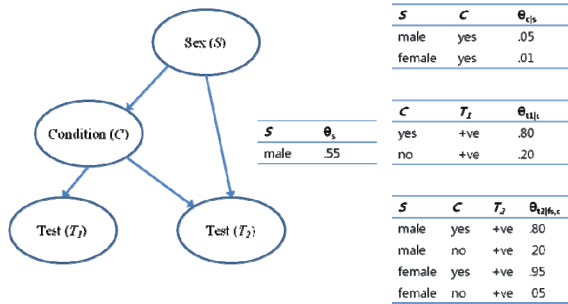


그림 3. 특정 인구, 질환 및 두 가지 검진 기법을 모델링하는 베이저안 망.  $C=yes$  는 해당 질환이 있음을 의미하고,  $T_i=+ve$  는 검진기법  $T_i$  의 결과가 양성임을 의미한다.

의 핵심적인 특성을 바로 유도해낼 수 있다. 시간  $t$  일 때 시스템의 상태를 알면, 이후의 시간  $s>t$  에서의 상태와 출력 값은 이전 시간  $s<t$  의 상태 및 출력 값과 독립이다.

그림 2b 의 네트워크는 동적 베이저안 망 (DBN, dynamic Bayesian network)으로 알려진 모델의 가장 단순한 형태라는 점도 알아두자. HMM 모델을 확장한 다양한 모델은 보다 복잡한 구조를 가지는 DBN 의 예로 볼 수 있다. 보통 기존 모델 확장 시 반드시 관련 추론 알고리즘의 집합을 병행하여 개발해야 하지만, HMM 을 베이저안 망으로 해석함에 따라 베이저안 망에서 가용한 알고리즘을 바로 적용할 수 있다는 장점이 있다.

### 3. 베이저안 망 구축하기

#### 3.1. 알고리즘

베이저안 망을 구축하는 방법은 크게 세 가지가 있다. 첫 번째는 지극히 주관적인 기

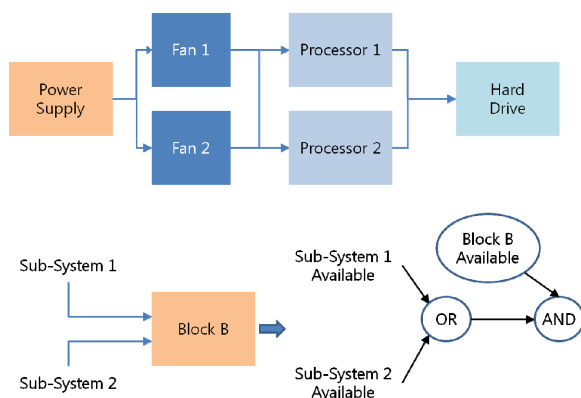


그림 4. 신뢰성 블록 다이어그램(위) 및 블록을 베이저안 망 조각으로 변환하는 체계적인 기법(아래)

법으로서, 기존에 알고 있는 지식, 특히 인과적 요인에 대한 개념을 베이저안 망에 반영하는 방법이다. 그림 1 에 표현한 베이저안 망이 바로 이 기법으로 구축한 예이다. 그림 3 에 구축한 또 다른 예제는 동일한 기법을 적용하되 파라미터 값은 인구통계학 및 검진기법의 명세와 같은 보다 공식적인 데이터에서 얻을 수 있다. 이 망에 따르면, 남성 55%, 여성 45%로 구성된 인구집단이 있고 그 중 의학적 질환  $C$  에 걸릴 가능성이 남성의 경우 더 크다. 이 질환을 검사할 수 있는 두 검진기법  $T_1$  과  $T_2$  가 알려져 있으며,  $T_2$  는 여성의 경우 더 효과적인 진단을 하고, 남성에게 대해서는 두 검사의 성능이 동일하다.

베이저안 망을 구축하는 두 번째 기법은 잘 정리되어 있는 다른 형태의 지식을 기반으로 자동 구축하는 것이다. 시스템에 대한 신뢰성 분석이나 진단이 필요한 수많은 응용사례에서 규정화된 시스템 설계를 기반으로 자동으로 베이저안 망을 합성할 수 있다. 그림 4 는 신뢰성 분석에 사용되는 신뢰성 블록 다이어그램(RBD, relevance block diagram)의 예이다. RBD 는 시스템 구성 요소 및 요소의 가용성 간 연관관계를 표현한다. 예를 들면, 프로세서 1 번은 두 팬 중 하나라도 동작하는 것이 필요하며, 각 팬은 전원을 필요로 한다. 분석 목표는 각 요소의 신뢰도를 알고 있을 때 시스템 전체의 신뢰도, 즉 가용확률을 계산하는 것이다. 그림 4 의 하단에는 RBD 의 각 블록을 체계적으로 베이저안 망 구성 요소로 변환하는 방법이 표현되어 있다. 이 결과 그림 5 와 같은 베이저안 망을 구축할 수 있다. 이 그림에서 CPT 는 각 개별 요소의 신뢰성과 변환기법의 의미적 요소를 기반으로 완벽히 구성할 수 있다.

세 번째 방법은 데이터를 기반으로 베이저안 망을 학습하는 기법이다. 예를 들어 다음과 같은 데이터 표에서 그림 3 과 같은 베이저안 망을 학습한다고 하자.

Sex $S$	Condition $C$	Test $T_1$	Test $T_2$
male	no	?	-ve
male	?	-ve	+ve
female	yes	+ve	?
⋮	⋮	⋮	⋮

이러한 데이터를 이용하여, 구조가 결정

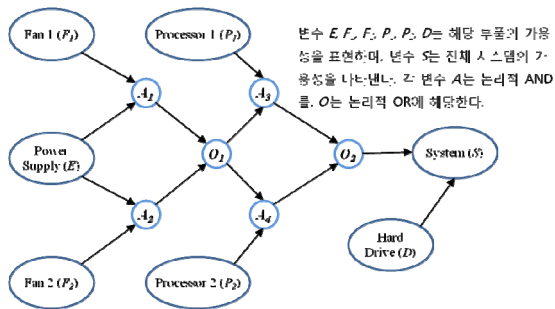


그림 5. 신뢰성 블록 다이어그램에서 자동으로 생성한 베이지안 망

된 베이지안 망의 파라미터만 학습할 수도 있고, 구조와 파라미터를 모두 학습할 수도 있다. 파라미터만 학습하는 것이 훨씬 용이한 작업이며, 데이터의 모든 값이 빠짐없이 알려져 있는 경우 보다 효율적인 학습이 가능하다.

학습은 귀납적 과정이므로, 이 학습 과정을 안내할 귀납적 원리가 필요하다. 두 가지 핵심 원리는 최대 우도 및 베이지안 방식이다(참고문헌 [2-5] 참조). 최대 우도 방식은 주어진 데이터가 관측될 확률을 최대로 하는 베이지안 망을 선호하며, 반면에 베이지안 방식은 우도 원리에 더하여 선호되는 베이지안 망 정보를 반영하는 사전 정보를 추가로 고려한다. 망의 파라미터만을 학습하는 문제에 대해 베이지안 방식으로는 각 파라미터 별로 가능한 값에 대한 사전확률분포를 부여하며, 데이터를 관측한 결과와 결합하면 파라미터에 대한 사후확률을 구할 수 있다. 이 사후확률을 파라미터 값 하나를 결정하는 데 사용할 수 있다. 또는 파라미터 값을 고정시키는 대신 질의에 대한 답을 계산하는 데에 사후확률에 따라 정해질 수 있는 모든 파라미터를 기준으로 평균을 구하는 방식도 적용할 수 있다.

망 학습 단계에서 “베이지안 망”이라는 명칭이 반드시 베이지안 방식을 적용해야 됨을 의미하지는 않는다. 베이지안 망이라는 명칭은 Judea Pearl이 세 가지 주요 특성을 강조하기 위해 부여하였다[6]. 첫째는 베이지안 망 구성에 사용되는 정보는 종종 주관적인 특성을 가진다는 점이고, 둘째는 추론 단계에서 베이즈 정리가 적용되는 점이며, 셋째는 망에 기반하여 증거를 바탕으로 한 추론뿐만 아니라 인과적 추론도 가능하다는 점이다.

지금까지 살펴본 학습 방식으로 구성된 베이지안 망은 풀고자 했던 작업과 상관없이

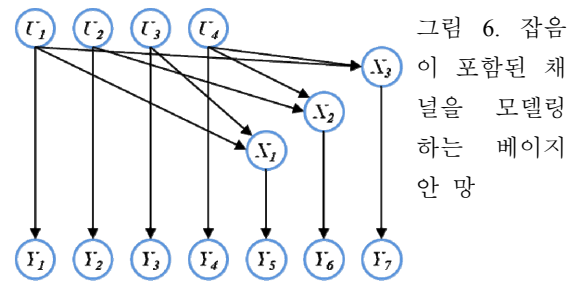


그림 6. 잡음이 포함된 채널을 모델링하는 베이지안 망

그 자체로도 의미가 있다. 예를 들어, 여러 질병 관련 증상을 모델링하는 망을 고려해보면, 망은 ‘진단’용으로 쓸 수도 있고(관측된 증상을 바탕으로 가장 가능성이 높은 질병을 추론), ‘예측’용으로도 적용할 수 있다(특정 질병에 대해 가장 가능성이 높은 증상을 추론). 한 가지 예로서, 진단을 위한 용도만을 고려하는 경우 학습한 망에서 진단 성능을 최적화하는 보다 특화된 귀납적 원리를 활용할 수 있다. 이러한 경우 기계학습 분야의 용어로는 ‘구분형(discriminative) 모델’을 학습한다고 표현한다. 이와 구분되는 방식은 ‘생성형(generative) 모델’이며, 이 모델은 특정 작업 수행 능력과는 관계 없이 주어진 데이터 집합을 생성할 수 있는 능력을 바탕으로 성능을 평가한다.

### 3.2. 규정화된 베이지안 망

다양한 분야의 대표적 문제에 대한 모델로서 정규화된 베이지안 망 모델이 제시되고 있다.

잡음이 존재하는 채널을 통해 정보를 전달하는 문제에서는 채널의 출력 값을 기준으로 가장 원본에 근사한 메시지를 계산하는 것을 목표로 하며, 그림 6에 예로 든 베이지안 망과 같은 규정화된 모델이 알려져 있다.

컴퓨터 비전 분야의 많은 문제에 또 다른 중



그림 7. 왼쪽부터 입력 그림, 베이지안 망 추론을 이용해 복원한 그림 및 원본 그림



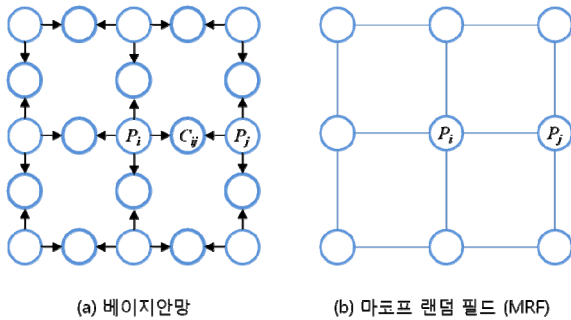


그림 8. 저수준 비전 문제에 적용 가능한 두 형태의 그래프 모델: 베이지안 망 및 마코프 랜덤 필드

류의 규정화된 베이지안 망이 적용되고 있다. 그림 7에 표현한 이미지 복원 문제의 경우, 손상된 이미지를 베이지안 망에 입력하고 추론을 통해 원본을 복원하였다. 이러한 문제에는 그림 8에서와 같이 베이지안 망과는 다른 그래프 모델인 MRF(Markov random field)를 기반으로 모델을 구성할 수 있다. 베이지안 망 모델의 경우 각 픽셀  $P$  사이에 스무딩 제한 조건으로서 픽셀의 자식 노드  $C$ 를 추가하고 픽셀의 채도에 대한 사전 확률을 정의하여 모델을 구성할 수 있으며, 이러한 모델은 스테레오 비전, 사진몽타주, 영역 구분과 같은 “픽셀 분류” 문제에서 깔끔한 추론 기법 유도 및 훌륭한 성능을 보인다.

최근 수 년간에는 규정화된 베이지안 망을 이용한 문서 분석과 같이 새로운 맥락의 응용이 등장하고 있다. 이 망의 다수는 토픽 모델을 기반으로 하는데, 토픽 모델은 문서가 숨겨진 토픽에 의해 생성된다고 보며 단어, 문서 및 감춰진 토픽 간의 연관성에 대한 추론 프레임워크를 갖추고 있다. 토픽 모델은 전자우편, 과학논문의 초록, 대규모 신문 문서와 같은 다양한 종류의 문서에 응용되고 있으며, 베이지안 망에서의 추론을 통해 문서간 유사도를 측정하고, 새로 쓰는 토픽을 찾거나 문서의 의미적 구조를 분석하는 등의 작업에 적용할 수 있다.

## 4. 베이지안 망을 이용한 추론

다른 모델링 언어와 비슷하게 베이지안 망의 가치는 망을 통해 해결 가능한 질의의 종류에 밀접하게 연관되어 있다.

그림 3을 예로 들면 다음과 같은 질의가 가능하다 - 두 검사에서 모두 양성으로 판별

된 남성이 실제 해당 질환을 보일 확률은? 인구 내의 어느 그룹이 두 검사에 대해 음성으로 판별될 가능성이 큰가? 그림 5에서 시스템의 전체적 신뢰도는? 시스템이 동작하지 않는 상태일 때 두 팬은 어떤 상태에 있을 확률이 높은가? 시스템 전체의 신뢰도를 5% 올리려면 어떤 부품 하나를 교체하면 되는가? 그림 6에서 채널 출력 값이 1001100이 되도록 하는 가장 가능성이 높은 채널 입력 값은?

이러한 질문에 대한 체계적인 해법은 세 가지의 규정화된 베이지안 망 질의를 이용하여 제시할 수 있다. 베이지안 망을 활용하는 주된 장점은 해결하고자 하는 문제를 위해 특화된 알고리즘을 만드는 대신, 이제 소개할 지정된 질의에 대한 기준에 개발된 베이지안 망 알고리즘을 쓸 수 있다는 점이다.

### 4.1. 증거 관측 확률

이 질의는 확률값  $Pr(\mathbf{e})$ 를 관측한다.  $\mathbf{e}$ 는 베이지안 망 내의 일부 변수집합  $\mathbf{E}$ 에 대해 값을 부여한 결과이며, 변수값 할당 또는 증거라고 부른다. 예를 들어 그림 3에서 한 개인이 두 검사에 대해 모두 양성인 나올 확률의 경우 증거 관측 확률은 질의  $Pr(T_1=+ve, T_2=+ve)$ 를 통해 얻을 수 있다. 동일한 질의를 그림 5에 적용하여 시스템 전체의 신뢰도  $Pr(S=avail)$ 을 계산할 수 있다. 이 질의에 대한 결정문제는  $PP$ -완전임이 알려져 있다.

### 4.2. 가장 가능성 높은 설명(MPE)

베이지안 망 내 변수집합  $\mathbf{E}$ 에 대해 변수값 할당  $\mathbf{e}$ 가 주어진 경우, 이 형태의 질의는  $\mathbf{E}$ 의 여집합  $\mathbf{Q}$ 의 변수값 할당  $\mathbf{q}$  중에서 확률  $Pr(\mathbf{q}|\mathbf{e})$ 를 최대화하는 조합을 판별한다. 그림 3에서  $\mathbf{e}$ 는  $T_1=-ve, T_2=-ve$ 로 설정하고  $\mathbf{Q}=\{S, C\}$ 로 둔 후 MPE 질의를 이용하여 두 검사 결과가 모두 음성이 될 가장 확률이 높은 성별+질환 조합을 찾을 수 있다. 그림 7과 그림 8의 이미지 복원을 위한 MPE 질의는  $\mathbf{e}$ 를 모든  $ij$ 에 대해  $C_{ij}=true$ 로 두고  $\mathbf{Q}$ 를 모든  $i$ 에 대해  $P_i$ 로 두면 된다. MPE의 결정문제는  $NP$ -완전임이 알려져 있으며, 복잡도 이론에서의 표준적 가정 하에서 증거관측확률 질의보다 용이한 문제이다.

### 4.3. 사후확률 최대 가설(MAP)

베이지안 망 내 변수집합  $E$  에 대해 변수값 할당  $e$  가 주어진 경우, 이 질의는  $E$  의 여집합 중 부분집합  $Q$  에 대한 변수값 할당  $q$  중에서 확률  $Pr(q|e)$ 를 최대화하는 조합을 판별한다. MAP 은 MPE 보다 어려운 문제이며 그 결정문제는  $NP^{PP}$ -완전임이 알려져 있다.

이들 규정한 질의를 이용하여 보다 복잡한 질의를 구현할 수도 있다. 민감도분석에 이러한 질의가 필요한데, 이는 베이지안 망에서 망의 파라미터가 섭동에 대한 튜닝한 정도를 점검하는 과정이다. 민감도 분석은 전문가가 설정한 전체적인 제한조건을 맞추기 위해 자동으로 파라미터를 재조정하는 과정에도 적용할 수 있다. 예를 들어 그림 5 의 망 전체 신뢰도를 계산한 결과 95%로 측정되었다고 하자. 신뢰도를 99% 이상으로 향상시키기 위해 살펴보아야 할 부품 및 목표달성에 필요한 이 부품의 신뢰도를 민감도 분석을 통해 파악할 수 있다.

## 5. 베이지안 망 규모 vs. 계산

베이지안 망에 대한 추론 알고리즘은 크게 정확계산 알고리즘과 근사 알고리즘의 두 종류로 구분된다. 정확계산 알고리즘은 예측의 정확성을 줄이는 그 어떤 타협도 않기 때문에 계산량이 매우 많은 경향이 있다. 이 계열 알고리즘의 계산 복잡도는 망의 트리폭(treewidth)에 대해서만 기하급수적으로 증가하는데, 망 내부적으로 많이 연결될수록 추론의 복잡도가 증가한다. 그림 8 의 예에서 그림이  $n \times n$  픽셀로 구성되었다고 가정하면 트리폭은  $n$  이 되는데, 적절한 크기의 이미지라 해도 트리폭 알고리즘을 적용하기에는 복잡도가 너무 높다.

그 동안 그래프 구조를 기반으로 제거와 조건부에 기반한 정확도 알고리즘이 연구되었는데, 망의 파라미터에서 발견되는 추가의 구조를 반영하면 보다 효율적인 추론이 가능하다는 것이 밝혀졌다. 파라미터 수준의 구조는 결정성(0 또는 1 값을 가지는 파라미터) 및 내용 기반 독립성( $d$ -분리로는 파악되지 않는 독립성)이 있으며, 최근 트리폭이 매우 큰 망에 대해서도 정확한 추론이 가능한 알고리즘이 개발되어 유전적 연결성 분석이나 연계성 모델에서 생성되는 망에 적용되었다.

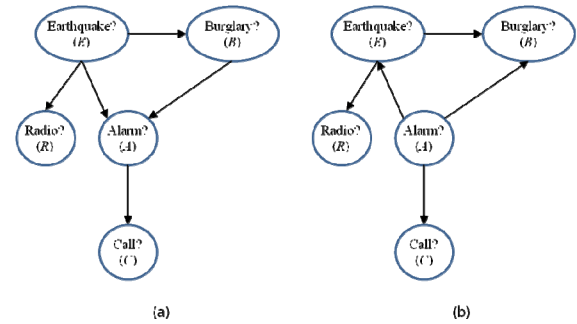


그림 9. 동일한 조건부 독립 관계를 표현하는 두 베이지안 망

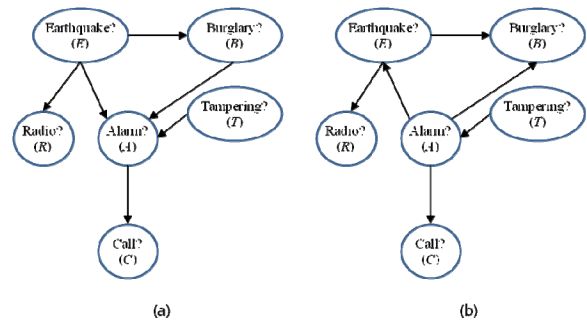


그림 10. 간섭을 고려하여 베이지안 망 확장하기

근사 알고리즘은 1990 년대 중반 이후 집중적으로 개발되었으며 일반적으로 트리폭에 무관한 계산 속도를 보인다. 현재 트리폭이 크되 파라미터 구조가 충분하지 않은 경우 근사 알고리즘이 유일한 선택이며, 저수준의 비전 응용문제는 대개 이러한 특성이 있다. 유력한 근사 알고리즘 계열로는 순환적 신뢰전파(loopy belief propagation) 기법, 확률론적 샘플링, 변분적(variational) 기법 등이 있다.

## 6. 인과성에 대한 추가 조명

베이지안 망에서 가장 흥미로운 부분은 인과성을 체계적으로 연구하는 과정에서 담당하는 역할이다. 그림 9 에는 동일한 변수 집합에 대한 두 가지의 베이지안 망 구조가 표현되어 있다.  $D$ -분리를 통해 두 구조가 동일한 조건부 독립성을 표현한다는 것을 확인할 수 있다. 적절히 CPT 를 부여하면 동일한 확률 분포를 표현하도록 만드는 것도 가능하며, 이 경우 두 베이지안 망은 기본적 표현 수준에서 동등하다. 그러나, 그림 9a 가 일반적인 인과성 관계에 대한 이해와 일치하는 반면 그림 9b 에는  $A \rightarrow E$ ,  $A \rightarrow B$  와 같이 상식에 반하는 연결이 포함되어 있다. 이러한 차이는 중요한

가? 한 망에서 표현되는 것이 다른 망에서는 얻을 수 없을 만한 정보가 있는가? 인과적 베이저안 망 관련 연구에 따르면 이에 대한 답은 ‘그렇다’이다.

이 연구에 따르면, 관측 결과가 아닌 ‘간섭’을 기반으로 한 확률변화 계산은 인과적 망을 이용해서만 처리할 수 있다고 한다. 그림 9의 예에서 누군가가 경보기를 허락 없이 개봉하여 경보기가 꺼진 경우를 생각해 보자. 이러한 상황은 간섭이며, 경보기가 꺼졌다는 것은 알지만 그 이유는 모르는 일반적인 관측 상황과는 다르다. 인과적 베이저안 망에서는 간섭을 그림 10에서와 같이 경보기 변수의 직접적 요인으로 추가한다. 이러한 부분적 수정을 그림 9b에 적용하여 그림 10b와 같은 망이 생성되면, 경보기가 꺼진 것을 관측하였을 때, 꺼진 이유가 불법개봉 때문인지의 여부는 도둑이 들었는지 또는 지진이 발생했는지 여부와는 하등 상관이 없다는 결론이 나게 되며(d-분리를 이용), 이러한 독립성 추론은 직관에 반하는 결과이다. 이 예제에서 살펴볼 수 있듯이, 인과성이 부적절하게 반영된 구조의 베이저안 망은 간섭에 연관된 질문에는 답할 수 없다.

인과성 구조를 이용하여 반(反)사실성과 같은 보다 복잡한 질문을 처리할 수 있다. 예를 들어, “그 환자가 약을 먹지 않았다면 생존하였을” 확률 계산에는 간섭요소에 대한 추론이 필수적이며, 때로는 표준적인 인과성 베이저안 망 이상의 기능적 정보가 필요할 수도 있다. 다른 질문의 예로 직접적 요인과 간접적 요인 구분 및 인과관계를 야기하는 필요조건과 충분조건을 결정하는 질의가 있다.

## 7. 베이저안 망의 확장

베이저안 망은 확률 그래프 모델(PGM, probabilistic graphical models)이라 부르는 보다 넓은 범위의 모델 중 한 종류로 볼 수 있다. 다른 모델로는 마코프 망 및 연쇄 그래프가 있으며, 마코프 망은 방향성이 없는 구조, 연쇄 그래프는 방향성의 유무가 혼합된 구조를 가진다. 각 모델은 확률 분포를 각각 다른 방식으로 압축 표현하는데, 확률에 기반한 의미 파악에 있어서 베이저안 망이 다른 모델에 비해 명시적이다. 그림 8b는 마코프 망의 한 형

태로서 비전 분야 응용에 주로 쓰이는 마코프 랜덤필드의 예이다.

베이저안 망을 확장한 형태의 관련 기법으로는 확률적 믿음에 더불어 효용과 결정을 표현하는 노드를 추가한 형태의 영향성 도표(influence diagram)[7], 규칙 및 반복을 토대로한 템플릿 기반 베이저안 망, 일차논리식 기반의 언어를 결합한 베이저안 망 등이 있다.

## 8. 남겨진 과제 및 분석자 결론

베이저안 망은 불확실성이 존재하는 상황에서 언제든지 쓸 수 있는 모델링 및 추론 기법으로 자리매김하고 있다. 도구로서의 베이저안 망이 발전하기 위해 필요한 것은 표현 능력을 확대하되, 베이저안 망의 성공에 필수 요소인 모듈 기반의 표현방식, 명료하게 표현되는 그래프 구조, 추론의 효율성 유지이다. 추론 알고리즘의 발전 및 특성 파악도 필수적이다.

지금까지 베이저안 망의 개념, 특성 및 관련 알고리즘과 응용에 대해 전반적으로 살펴 보았다. 앞으로 다양한 응용 문제의 해결로서 많은 사람들이 베이저안 망의 유용성을 이해하고 활용하기를 기대하며 본 고를 마친다.

## 참고 문헌

1. A. Darwiche, *Bayesian Networks*, *Communications of the ACM*, December, 2010.
2. R. Cowell, A. Dawid, S. Lauritzen, and D. Spiegelhalter, *Probabilistic Networks and Expert Systems*, Springer, 1999.
3. A. Darwiche, *Modeling and Reasoning with Bayesian Networks*, Cambridge University Press, 2009.
4. D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, Cambridge, MA, 2009.
5. R. Neapolitan, *Learning Bayesian Networks*, Prentice Hall, Englewood, NJ, 2004.
6. J. Pearl, Bayesian networks: A model of self-activated memory for evidential reasoning, In *Proceedings of the Cognitive Science Society*, 1985.
7. F. V. Jensen and T. D. Nielsen, *Bayesian Networks and Decision Graphs*, Springer, 2007.