

Dempster-Shafer 이론을 이용한 징후 결합 방법

(An Evidence Fusion Method Based on the Dempster-Shafer Theory)

이 증 만* 김 진 형**
(Jong Man Lee), (Jin Hyung Kim)

요 약

Dempster-Shafer 이론은 징후 결합을 위한 수학적 방법을 제공하고 있다. 이 이론에서 각 가설의 신뢰도는 그 가설과 그의 부분집합들에 배당된 기본적 신뢰도 값의 합으로 정의된다. 그러나, 가설의 숫자가 많아지면 가설들의 부분집합의 수가 기하급수적으로 증가하므로, 징후 결합시 모든 부분집합을 고려할 수 없다. 따라서 이러한 복잡도를 낮추기 위하여 여러가지 방법이 제안되었다. 이 논문에서는, 고려하는 부분집합을 주어진 가설과 그의 보수 가설들로 제한하는 새로운 근사방법을 제안하였다. 즉, Dempster의 징후결합방법에 의해 생성되는 새로운 가설이 주어진 가설집합과 그의 여집합들에 포함되어 있지 않으면, 이 가설에 대한 신뢰도 값을 가설집합들과 그의 여집합들내에서 주어진 가설의 가장 가까운 상위 가설에 배당하였다. 이 논문에서 제안된 방법을 이용함으로써, 근사오차를 줄일 수 있고 신뢰 구간의 개념도 유지할 수 있는 장점을 얻을 수 있음을 보였다. 결론 데이터를 이용하여 제안한 방법의 성능을 평가하였다.

ABSTRACT

The Dempster-Shafer theory offers a mathematically solid framework for fusing evidence. According to the theory, the belief on a hypothesis is computed by aggregating all the basic belief, called bpa, committed to itself and its subsets. However, since the set of subsets grows exponentially with the number of elements, not all subsets can be counted for evidence combination. Therefore, various approximation schemes are proposed to restrict the subsets being considered. In this paper, a new approximation scheme is proposed which restricts the subsets as the given hypotheses and their complements. That is, if a new hypothesis generated by the Dempster's orthogonal sum is not contained in the given hypotheses and their complements, its belief is assigned to the nearest supersets among the hypotheses and their complements. The proposed scheme yields much closer approximation and still maintains the concept of the belief interval. The performance of the proposed scheme is verified through an evaluation with a sample data set.

* 정 회 원 . 한국과학기술원 전산학과

** 종신회원 한국과학기술원 전산학과 조교수

접수일자 1987 9 14

1. 서 론

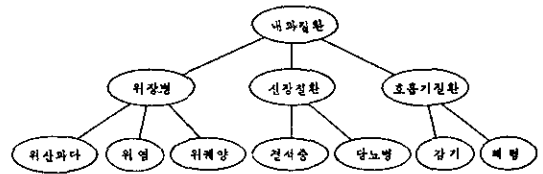
전문가 시스템이 여러 분야에 널리 사용되면서 주어진 자료가 비록 불완전하고 불충분하더라도 이 자료를 이용하여 인간 전문가가 내리는 판단과 비슷한 결과를 얻을 수 있도록 하기 위한 불확정성의 문제가 더욱 중요시 되고 있다. 인공지능 시스템에서 불확정성의 문제 해결을 위하여 사용되는 방법에는 여러가지가 있다. 첫째, 확률에 기초를 두고 있는 베이스 이론[1]은 오랜동안 연구되었고 그 효용성이 입증되었으나, 이 방법을 적용하기 위해서는 얻기 힘든 많은 자료가 요구되는 단점이 있다. 다른 하나의 방법은 내과 진료용 전문가 시스템인 MYCIN에 이용된 불확정 지수(Certainty Factor) 방법[2,3]으로서, 개념이 비교적 간단하여 사용시 효율성은 좋으나 수학적 기반이 없는 단점이 있다 베이스 이론 대신에 사용할 수 있는 Dempster-Shafer(D-S) 방법[4,5]은 최근에 많은 연구가 되고 있는 방법으로서 불확정 지수 방법보다는 효율성이 나쁘며 고려되는 전체 영역내의 집합들이 서로 독립적인 성격을 가져야 하는 전체 조건이 필요하지만, 신뢰구간의 표현 및 고려되는 가설들에 대해 아직 모르는 부분에 대한 표현이 가능하다는 장점이 있다 이 D-S 방법은 주어진 징후들을 결합시키는 과정에서 기하급수적 복잡도를 가진다 이 복잡도를 줄이기 위해, Barnett[6]은 고려하는 문제 영역을 단일 가설만으로 제한하는 방법을 제안하였다. 따라서 복잡한 가설에 대한 신뢰도는 처리할 수 없는 단점을 가진다. Gordon과 Shortliffe[7]는 Dempster의 징후 결합 방법 사용시 생성되는 가설에 해당될 신뢰도 값을, 주어진 가설 집합내에서 가설의 가장 가까운 상위 가설에 근사시켜 배당하는 방법을 제안하였다. 이 방법을 이용함으로써 복잡도는 줄일 수 있지만, 두가지 단점이 제기되었다. 첫째, 신뢰도 값이 상위 가설에 근사되어 배당되므로써 근사 오차가 발생한다 두번째 단점으로서, 계산시 고려하는 가설 집합을 확증 가설들로만 제한하여, 부가적인 근사 계산을 하지 않으면 각 가설에 대한 불확증 신뢰도를 알 수 없다 따라서, 징후 결합 후에 D-S이론의 가장 큰 장점인 신뢰구간의 개념을 유지할 수 없다는 단점을 가진다 이 논문에서는, 징후 결합 계산시 근사 오차를 줄

일 수 있고 신뢰구간의 개념을 유지할 수 있는 새로운 근사 방법을 제안하였다. 즉, 고려하는 가설집합을 주어진 가설집합과 이들의 여집합들로 확장하였다. 징후 결합시 생성되는 새로운 가설에 해당되는 신뢰도값을 위의 확장된 가설집합 중에서 생성된 가설의 가장 가까운 상위 가설에 배당하는 방법을 제시하였다. 이러한 방법을 이용함으로써, 징후 결합시 Gordon과 Shortliffe의 방법보다 근사 오차를 줄일 수 있고, 신뢰구간의 개념을 유지할 수 있는 장점을 얻을 수 있다

이 논문의 구성은 다음과 같다 제 2장에서는 징후 결합을 위한 D-S 이론의 개요를 설명하였고, 제 3장에서는 D-S 이론의 복잡도를 줄이기 위해 제안된 기존의 두가지 방법을 설명하였다. 제 4장에서는 이 논문에서 제안한 방법에 대해 설명하였고, 제 5장에서는 간단한 견본 데이터를 이용하여 제안한 방법의 성능을 평가하였으며, 이 논문의 결론을 제 6장에 기술하였다

2. 분할가능 단조 직교 다각형에 대한 알고리즘

전문가 시스템의 추론기능은 고려되는 각 가설들에 대해 주어진 자료를 종합하여, 가능성 있는 각 가설에 대해 그 가능성의 정도를 나타내 주는 과정으로 볼 수 있다. 본 논문의 이해를 돕기 위하여 내과의사가 진료하는 질병 중에서 위장병과 호흡기 질환 및 신장 질환을 예로 들어 설명하겠다. 위장병은 결석증과 당뇨병으로 각각 세분되며, (그림1)이 이들 질병 사이의 연관관계를 설명하고 있다



(그림 1) 내과 질환의 계층 구조

D-S 이론에서 위의 7가지 질병, {위산과다, 위염, 위궤양, 감기, 폐렴, 결석증, 당뇨병}을 고려되는 전체 구조(Frame of Discernment)라 하며 Θ 로 표시

한다 이 θ 내의 각 항목들은 상호 배반적이며(Mutually Exclusive) 전 영역을 포함하는(Exhaustive) 성격을 가지며, 상위 계층의 질병은 하위 계층의 질병들을 포함한다 즉, 한가지 질병은 하위 계층의 다른 질병들의 집합으로 구성된다. 위(그림 1)의 11 가지 각 질병들에 대해 각각 확증 자료와 불확증 자료가 제공될 수 있다

θ 의 각 부분집합들에 대한 고유의 자료는 기본 확률 배당 함수(basic probability assignment, bpa)로써 나타내며, $m(\cdot)$ 으로 표시한다. θ 의 한 부분집합 A에 대한 $m(A)$ 의 값은 A 자체에만 배당된 확신도 값을 나타내며, 이 값은 A를 구성하는 다른 부분집합의 확신도 값으로 세분되지 않는다 $m(\theta)$ 는 θ 에 배당되는 확신도 값으로서 θ 의 부분집합들에 배당되고 아직 남아있는 확신도 값을 나타낸다 각 질병에 대한 신뢰도(Belief) 함수는 $Bel(\cdot)$ 로 표시되며 임의의 부분집합 A에 대한 신뢰도 값은 A의 모든 부분집합들의 기본 확률 배당 함수값을 합한 값으로 정의된다. 즉, $Bel(A)$ 는 가설 A에 대한 총신뢰도의 합을 나타낸다. 이를 식으로 표현하면 다음과 같다

$$Bel(A) = \sum_{B \subset A} m(B) \dots\dots\dots (1)$$

각 가설에 주어지는 확증, 불확증 자료들은 상호 결합되는 과정이 필요하다 이를 위해 D-S 이론에서는 θ 의 부분집합이면서 서로 다른 관찰에 의해 수집된 징후 값들을 결합시키는 방법으로써 \oplus 로 표시되는 징후 결합 방법을 제공하고 있다 m_1, m_2 가 서로 다른 기본 확률 배당 함수일 때, 이들이 결합된 새로운 결과는 다음 식으로 나타낼 수 있다

$$m_1 \oplus m_2(A) = \frac{\sum_{A \cap B_j = A} m_1(A_1)m_2(B_j)}{1 - \sum_{A \cap B_j = \phi} m_1(A_1)m_2(B_j)} \dots\dots (2)$$

즉, m_1 과 m_2 가 결합된 후의 가설 A에 대한 bpa 값은, m_1 과 m_2 의 각 부분집합 X, Y중에서, X와 Y의 공통집합이 A가 되는 값들을 모은 것이다. 이때, 두 부분집합의 공통집합이 공집합이 되면 이에 지정된 값을 뺀 나머지의 합이 1이 되도록 재조정한다 주어진 가설 A에 대한 신뢰도 $Bel(A)$ 와, A의 보수 가설에 대한 신뢰도 $Bel(A^C)$ 를 이용하여

신뢰구간이 다음과 같이 정의된다

$$[Bel(A), 1 - Bel(A^C)]$$

신뢰구간의 폭은 가설에 대한 불확실성, 즉 아직 알려지지 않은 정보의 양으로 해석할 수 있다. 신뢰구간으로써 아직 모르는 정보의 양을 나타낼 수 있다는 점이 베이스 확률 이론에 비해 D-S 이론이 가지는 가장 큰 장점이라 할 수 있다.

3. D-S 방법의 적용 예

3.1 단일 요소로만 구성된 문제에서의 적용

D-S 이론은 각 가설에 대한 불확증 신뢰도와 신뢰구간의 표현이 가능한 장점이 있다 Dempster의 징후 결합 방법 적용시 θ 의 모든 부분집합이 생성되므로, θ 내 항목 수의 증가에 따라 계산의 복잡도가 기하급수적으로 증가하게 된다. θ 의 모든 가능한 부분집합 중에서 단일 요소로 구성된 집합들에 대한 확증 자료와 불확증 자료만을 고려하여 복잡도를 낮추는 방법이 Barnett[6]에 의해 제안되었다. 이 방법은 다음과 같이 3단계로 나누어진다

1 단계 단일 요소로만 구성된 각 가설에 대해 확증 징후와 불확증 징후를 각각 결합시킨다

2 단계 각 가설별로 1 단계에서 구한 확증 bpa를 하나의 bpa로 결합시킨다 이 단계가 끝나면 가설 갯수 만큼의 bpa가 만들어진다.

3 단계 위의 단계 2에서 만들어진 가설 갯수 만큼의 bpa를 하나의 Belief 함수로 결합시킨다.

이 방법을 이용하므로써 선형의 복잡도로 주어진 징후들을 결합할 수 있으며, 각 가설 A에 대한 $Bel(A)$ 와 $Bel(A^C)$ 를 알 수 있어 신뢰구간을 구할 수 있는 장점이 있다. 그러나, 이 방법은 단일 요소로 구성된 가설에 대해서만 적용할 수 있다는 단점을 가진다.

3.2 Gordon과 Shortliffe의 근사 방법

D-S 이론의 복잡도를 선형으로 줄이기 위하여 Gordon과 Shortliffe는 새로운 근사 방법을 제안하였다 즉, 징후 결합시 고려하는 부분집합의 수를 제

한하여 생성되는 모든 부분집합들을 이들에 근사시켰다 두 bpa를 결합시, 고려되고 있는 가설에 대응되는 부분집합들에 존재하지 않는 새로운 부분집합 A가 생성되면, 이에 배당될 신뢰도 값을 A에서 가장 가까운 상위 부분집합에 배당하는 방법을 제안하였다. 이 방법은 3 단계로 나누어 진다.

1 단계 각 가설에 대해 확증 징후와 불확증 징후를 각각 결합시킨다. 이 단계는 Barnett의 방법과 동일하다.

2 단계 모든 주어진 가설에 대한 확증 징후만을 결합시킨다. 이 단계가 끝나면 하나의 확증 bpa와 전체 가설 갯수 만큼의 불확증 bpa가 남는다. 확증 징후의 결합시에는 주어진 가설에 해당되는 부분집합만이 생성되므로 복잡도의 문제는 제기되지 않는다.

3 단계 2 단계에서 생산된 하나의 확증 bpa에, 아직 결합이 되지 않은 불확증 bpa들을 하나씩 다음 방법으로 결합한다

(1) 임의 가설에 배당된 불확증 징후를, 그 가설의 역에 해당되는 부분집합에 배당한다.

(2) 이 징후를 2 단계의 결과인 하나의 확증 bpa에 결합한다. 이때 생성되는 부분집합 A가 주어진 가설 집합에 있으면, A에 배당될 bpa 값을 해당 가설에 배당한다. 만약 A가 주어진 가설 집합에 없으면, 주어진 가설 집합 내에서 A와 가장 가까운 상위 부분집합에 A에 배당될 bpa 값을 배당한다.

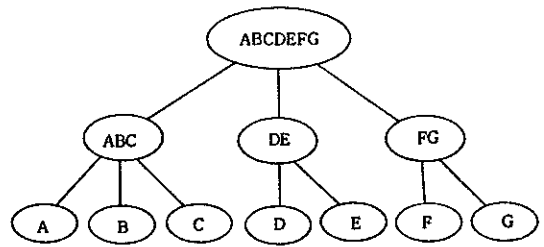
이러한 근사 방법을 이용하므로써, 선행의 복잡도로 모든 징후들을 결합시킬 수 있으나, 다음 두가지 단점이 있다. 첫째, 불확증 징후 결합시 사용된 근사 방법에 의해 근사오차가 생긴다. 이 오차는 생성된 부분집합이 주어진 가설 집합에 없는 경우 그와 가장 가까운 상위 가설에 bpa 값이 배당되기 때문에 발생한다. 두번째 단점으로서, 징후 결합 후에 각 가설에 대한 불확증 징후를 알 수 없다. 즉, 불확증 징후를 저장할 수 있는 부분집합이 없으므로 부가적인 근사 제산을 해야 이 값을 알 수 있다. 따라서, D-S 이론의 가장 큰 장점인 각 가설에 대한 신뢰구간도 알 수 없게 된다

4. 여집합을 고려한 징후 결합 방법

앞 장에서는 D-S 이론의 복잡도를 줄이기 위한 기

준의 두가지 방법을 알아보았다. Barnett의 방법은 적용할 수 있는 문제 영역을 단일 가설로만 구성된 경우로 제한하였다 Gordon과 Shortliffe의 방법에서는 징후 결합시 고려하는 부분집합의 수를 주어진 가설의 수로 제한하였다 이에 따라 근사 오차가 발생하였고, 신뢰구간을 유지할 수 없는 단점이 제기되었다. 이 장에서는 근사 오차를 줄일 수 있고, 신뢰구간의 표현이 가능한 새로운 징후 결합 방법을 제시하겠다.

먼저, 설명상의 편의를 위해 (그림 1)의 가설계층 구조를 기호적 표현으로 바꾸면 다음과 같다



(그림 2) 가설 계층 구조의 기호적 표현

계층 구조로 구성된 각 가설에 대해 확증 징후와 불확증 징후를 얻을 수 있다. 그러나 Gordon과 Shortliffe의 방법에서는 이 두가지 자료가 주어짐에도 불구하고, 징후 결합시 주어진 가설에 해당되는 부분집합들만 고려하여 확증 징후들의 값만 정확히 저장할 수 있었다. 불확증 징후는 그 가설의 여집합에 배당한 다음 이를 결합시킴으로써, 생산이 끝난 후 각 가설에 대한 불확증 신뢰도 값을 알 수 없는 단점이 제기되었다. 즉, Gordon과 Shortliffe 방법의 두번째 단점은 고려하는 부분집합의 수를 너무 제한하므로써 발생하였다. 주어진 징후에 대해서는 징후 결합 생산시 이를 정확히 표현하고 저장할 수 있는 부분집합을 고려하여야, 근사 오차를 줄일 수 있고 전체 징후의 결합 후에도 각 가설에 대한 불확증 신뢰도 및 신뢰구간의 유지가 가능해 진다. 이 논문에서는 징후 결합 생산시 고려하는 부분집합의 수를 주어진 가설에 해당되는 부분집합들과 이의 여집합들로 증가시켰다. 앞으로 주어진 가설에 해당되는 부분집합들과 이의 여집합들을 확장된 가설 집합 (Extended Hypothesis set)이라 하겠다. Dempster

의 방법을 이용한 징후 결합 계산시 생성되는 임의의 부분집합 A가 확장된 가설 집합에 없으면, A에 배당될 bpa 값을 확장된 가설 집합 내에서 A에서 가장 가까운 상위 부분집합에 배당한다. A에서 가장 가까운 상위 부분집합이 하나 이상 존재하면, A에 배당될 bpa 값을 A의 상위 부분집합들에 균등하게 분배한다 (그림2)를 이용하여 제안된 알고리즘을 설명하겠다 T는 주어진 가설에 대응되는 부분집합들의 집합을 나타낸다 T^C는 T의 각 항목에 대한 여집합들의 집합을 나타낸다 (그림 2)에서 T는 {{A, B, C}, {D, E}, {F, G}, {A}, {B}, {C}, {D}, {E}, {F}, {G}} (이를 간단히 {ABC, DE, FG, A, B, C, D, E, F, G}으로 나타냄)이고 T^C는 {ABC^C, DE^C, FG^C, A^C, B^C, C^C, D^C, E^C, F^C, G^C}이다. 우리가 최종적으로 구하고자 하는 바는 T와 T^C에 대해 주어진 모든 징후들을 결합하는 것이다. 이것은 다음식으로 표현된다

$$m_{Y_1} \oplus m_{Y_2} \oplus \dots \oplus m_{Y_i} \dots \dots \dots (3)$$

여기에서 Y_i는 T와 T^C에 있는 부분집합이다. 위의 식 (3)은 다음식으로 나타낼 수 있다.

$$m_T \oplus m_T^C \oplus m_{Y_i} \dots \dots \dots (4)$$

여기에서

$$m_T = m_{X_1} \oplus m_{X_2} \oplus \dots \oplus m_{X_i}, X_i \in T \dots \dots \dots (5)$$

이고,

$$m_T^C = m_{X_1^C} \oplus m_{X_2^C} \oplus \dots \oplus m_{X_i^C}, X_i^C \in T^C \dots \dots (6)$$

이다 이때 m_T에 T^C내의 임의 부분집합을 결합시 확장된 가설 집합에 없는 새로운 부분집합이 생성된다. 정확한 계산을 위해서는 θ의 모든 부분집합인 2^θ가 고려되어, 생성되는 모든 부분집합에 배당되는 징후값을 저장하여야 한다. 그러나 이 방법은 복잡도가 높아 Gordon과 Shortliffe는 고려하는 부분집합을 T로 제한하였고, 이 논문에서는 이를 T와 T^C로 제한하였다 즉, Gordon과 Shortliffe의 방법은 m_θ를 m_T로 근사시킴에 비해, 이 논문에서는 이를 m_{TUT^C}로 근사시켰다. 제안된 징후 결합 방법의 자세한 내용은 다음과 같다

1 단계 T의 각 가설 X_i에 대해 확증 징후를 각각 결합시킨다. 모든 X_i에 대해, 확증 징후를 결합

하여 m_{X_i}를 만들고, 불확증 징후를 결합하여 m_{X_i}^C를 만든다. (그림 2)의 제층 구조에 대해 계산하면, m_{ABC}, m_{DE}, m_{FG}, m_A, m_B, m_C, m_D, m_E, m_F, m_G, m_{ABC}^C, m_{DE}^C, m_{FG}^C, m_A^C, m_B^C, m_C^C, m_D^C, m_E^C, m_F^C, m_G^C 등이 만들어진다.

2 단계 모든 주어진 가설들에 대해 식 (5)를 적용하여 확증 징후만을 결합시킨다. m_T은 T의 부분집합 또는 θ에만 값을 배당한다. 주어진 예제에서,

$$m_T = m_{ABC} \oplus m_{DE} \oplus m_{FG} \oplus m_A \oplus m_B \oplus m_C \oplus m_D \oplus m_E \oplus m_F \oplus m_G$$

이다 이 단계에서는 불확증 징후를 결합하지 않고 단계 3으로 미룬다 이제 남은 일은 불확증 징후를 결합시키는 m_T ⊕ m_{T^C}의 계산이다. 이미 언급했듯이 m_{T^C}의 계산시에 T나 T^C에 없는 새로운 부분집합에 bpa 값이 배당된다. 앞의 예제에서 m_T ⊕ m_{A^C}의 계산시 주어진 가설 집합에 없는 새로운 부분집합 {BC}에 bpa 값이 배당된다. 또한 m_T ⊕ m_{A^C}와 m_{B^C}의 결합시에도 주어진 가설 집합에 없는 새로운 부분집합 {AB}에 bpa 값이 배당된다. 이러한 부분집합들은 확장된 가설 집합내에 존재하지 않으므로 이들에 배당되는 bpa 값을 처리하기 위하여 다음의 근사 방법을 제시하였다.

3 단계 단계 2에서 구해진 각각의 확증 bpa에 불확증 bpa 들을 하나씩 차례로 다음 방법에 의해 결합시킨다. T^C에서 부분집합 X₁^C를 선택하여 m_T ⊕ m_{X₁}^C를 계산한다. 이때 T와 T^C에 없는 새로운 부분집합 A가 생성되면 이에 배당되는 bpa 값을 A에서 가장 가까운 상위 부분집합에 배당한다. A에서 가장 가까운 상위 부분집합이 하나 이상 존재하면, 상위 부분집합들에 균등히 나누어 분배한다. T^C에서 X₂^C를 선택하여 (m_T ⊕ m_{X₁}^C) ⊕ m_{X₂}^C를 계산한다. 이러한 방법으로 T^C내의 모든 X_i^C에 대한 bpa를 결합한다. 이 과정을 식으로 나타내면 다음과 같다.

$$m_{TUT^C} = m_T \oplus m_{X_1^C} \oplus m_{X_2^C} \oplus \dots \oplus m_{X_i^C}, X_i^C \in T^C \dots \dots \dots (7)$$

이 단계의 계산이 완료되면, 주어진 모든 징후가 결합된 단 하나의 bpa 분산을 얻게 된다. 이 방법에서 2^θ의 모든 부분집합을 고려하지 않고 T와 T^C만

을 고려하여 생성되는 모든 부분집합들을 이 확장된 가설 집합에 근사시키므로써 bpa 값이 상위 부분집합에 배당되는 경우가 존재한다.

주어진 예제에서, 부분집합 {CDEFG}에 배당되는 bpa 값은 {CDEFG}의 가장 가까운 상위 부분집합 {A^C}와 {B^C}에 균등히 나뉘어 배당된다. 그러나 Gordon과 Shortliffe의 방법에서는 이 값이 {CDEFG}의 가장 가까운 상위 부분집합인 {ABCDEFG}에 배당되었다.

5. 평 가

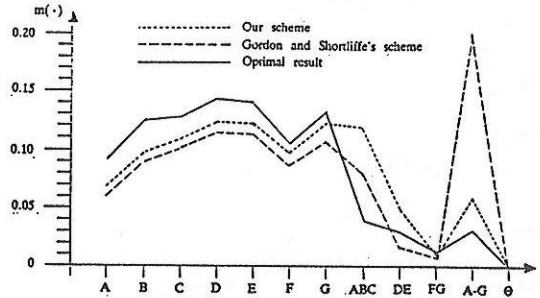
이 논문에서 제안한 방법은 Gordon과 Shortliffe의 방법보다 더 많은 부분집합들을 고려하기 때문에 부가적으로 복잡도가 높아진다. N이 주어진 가설의 수를 나타낼 때, N개 가설에 대해 각각 2번의 orthogonal sum 계산이 필요하며, 매 orthogonal sum 계산시 2N번의 곱셈이 필요하고, 각 곱셈의 결과를 배당할 가장 가까운 상위 가설을 찾는 계산량이 필요하다. 임의 가설의 가장 가까운 상위 가설을 찾는 계산량은, Gordon과 Shortliffe의 방법의 경우 $\log N$ 이지만 이 논문에서 제안한 방법에서는 N이다. 그러므로 총계산량은 각각 $O(N^2 \log N)$, $O(N^3)$ 이다. 비록 Gordon과 Shortliffe의 방법보다 약간의 기억 용량과 계산 시간을 더 필요로 하지만, 제안된 방법은 다음의 두가지 장점을 가진다.

첫째, 제안된 방법을 이용하므로써 근사 오차를 줄일 수 있다. 본 논문에서 제안한 방법을 이용함으로써 근사 오차를 줄일 수 있음을 부록에서 증명하였다. 본 논문에서 제안된 방법이 실제 정후 결합에 사용되는 예제로서, (그림 2)의 가설 계층 구조에 대해 다음 <표 1>의 견본 자료를 이용하여 제안된 방법의 성능을 평가하였다. 사용된 견본 자료는 다음과 같다.

<표 1> 견본자료

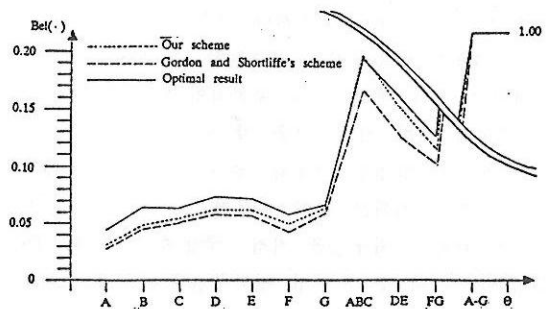
	confirmatory data	disconfirmatory data
A	0.50	0.62
B	0.60	0.61
C	0.70	0.72
D	0.55	0.54
E	0.58	0.58
ABC	0.52	0.64
DE	0.54	0.63
F	0.62	0.72
G	0.68	0.71
FG	0.51	0.64
ABCDEFG	0.52	0.58

다음 (그림 3)은 견본 자료를 이용하여 얻은 bpa 값이다.



(그림 3) 각 방법에 의한 bpa 값

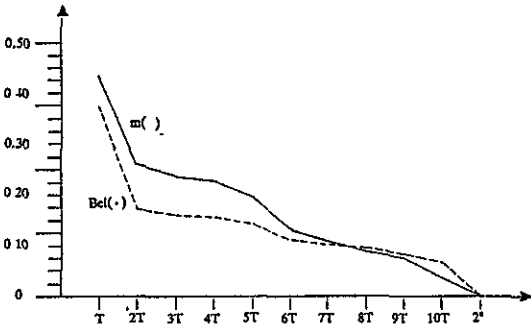
위의 (그림 3)에 나타나 있듯이 제안된 방법에 의한 결과가 Gordon과 Shortliffe 방법의 경우보다 대부분의 경우 오차가 적다. 그러나, {ABC} 등과 같은 중간 계층의 가설에 대해서는 제안된 방법의 경우 오히려 오차가 증가하였다. 정확한 계산에 의하면 단일 가설들에 배당될 값이, Gordon과 Shortliffe 방법에서는 근사 계산에 의해 보다 상위 계층에 많이 배당되었으나, 제안된 방법에서는 이 값의 상향화가 억제되어 중간 계층에 배당된 때문이다. 즉, 중간 계층에서는 그 자체의 오차와 하위 계층에서 발생된 오차가 축적되어 더욱 오차가 증가하였다. 그러나, 각 가설에 대한 신뢰도 값인 $Bel(\cdot)$ 을 계산시에는 이 두가지 오차가 서로 상쇄되어 중간 계층에서의 $Bel(\cdot)$ 오차가 현저히 감소되었다. 이와 같은 현상이 (그림 4)에 잘 나타나 있다.



(그림 4) 각 방법에 의한 Bel(.) 값

정후 결합 계산시 고려하는 부분집합 수의 변화에

다른 오차의 합이 (그림 5)에 나타나 있다. 여기에서 T 는 주어진 가설의 집합을 나타내며, n 은 θ 내 부분집합의 수를 나타낸다.



(그림 5) bpa 및 Bel(.) 오차의 경향

Gordon 과 Shortliffe 방법은 주어진 가설에 해당하는 부분집합들만 징후 계산시 고려하기 때문에 T 에 해당되며, 제안된 방법은 확장된 가설 집합(주어진 가설에 해당하는 부분집합 및 이들의 여집합)을 고려하기 때문에 $2T$ 에 해당된다. $2T$ 이상의 경우에는 확장된 가설 집합 2^{θ} 에서 임의로 T 개의 부분집합을 선택, 추가하여 계산에 이용하였다. 고려하는 부분집합의 수를 T 개에서 $2T$ 개로 증가시킴에 따라 오차가 급격히 감소하였다. 그러나, $3T$ 이상의 경우에는 증가하는 복잡도에 비해 오차의 감소율이 둔화되었다. 즉, 증가하는 복잡도와 오차의 감소 비율을 같이 고려한다면, 본 논문에서 제안한 방법인 $2T$ 인 경우가 가장 효율적이라 할 수 있다.

두번째 장점으로, 징후 결합 계산 후에도 각 가설에 대한 신뢰구간의 표현이 가능하다. Gordon 과 Shortliffe 의 방법에서는 징후결합 후 확증 가설에 대해서만 값이 배당되므로 부가적인 계산이 있어야 각 가설에 대한 신뢰구간을 알 수 있으나, 본 논문에서 제안한 방법을 이용하므로써 주어진 각 가설에 대한 불확증 신뢰도 값이 자동적으로 구해지므로 각 가설에 대한 신뢰구간을 직접 구할 수 있다. 즉, D-S 이론의 가장 큰 장점인 신뢰구간의 개념을 그대로 유지할 수 있는 장점을 가진다.

6. 결 론

이 논문에서는 Gordon 과 Shortliffe 의 방법보다 더

정확한 계산이 가능하고 D-S 이론의 장점을 유지할 수 있는 새로운 근사 방법을 제안하였다. 징후 결합 시, 주어진 가설에 대응되는 부분집합 뿐만 아니라 이들의 여집합들도 같이 고려하였다. Dempster 의 징후 결합 방법을 적용시 생성되는 모든 부분집합들을, 주어진 가설에 대응되는 부분집합 및 이들의 여집합에 근사시켰다. 이 방법은 Gordon 과 Shortliffe 방법보다 더 많은 부분집합을 고려하기 때문에, N 이 주어진 가설의 수일때 $O(N^3)$ 의 복잡도를 가진다. 제안된 방법을 이용하므로써, 두가지 장점을 얻을 수 있음을 보였다. 첫째 장점으로써, 이 방법은 Gordon 과 Shortliffe 의 방법보다 근사 오차를 줄일 수 있다. 기본 자료를 이용하여 그들의 방법에 의한 결과와 비교하여 오차를 줄일 수 있음을 보였다. 두번째 장점으로써 이 방법은 각 가설에 대한 불확증 신뢰도와 신뢰구간의 표현을 가능하게 한다. 각 가설의 불확증 신뢰도 값은 부가적인 계산없이 고려된 부분집합들 중에서 얻을 수 있다. 따라서 신뢰구간의 개념을 징후 결합 계산이 끝난후에도 유지할 수 있는 장점이 있다. 주어진 문제 영역의 가설들이 서로 배반적이고 전 영역을 포함하는 성격을 가지며 제층 구조로 구성될 수 있다면, 본 논문에서 제안한 방법을 이용하여 D-S 이론의 장점을 유지하면서 주어진 징후들을 효율적으로 결합할 수 있다.

부 록

T 및 T^C 가 이미 정의한 바와 같이 가설에 대응되는 가설들의 집합과 이들의 여집합을 각각 나타내며, θ 는 고려되는 전체 구조를, m_T 및 m_{TUT^C} 는 각각 식 (5), (6)과 같고, 평형 상수 K_T, K_{T^C} 는 다음과 같다.

$$K_T = K_{T^C} = 1 / [1 - m_A^C \sum_{\substack{X \in T \\ X \subseteq A}} m_T(X)]$$

본 논문에서 이용한 예제에서,

$$T = \{ \{A, B, C\}, \{D, E\}, \{F, G\}, \{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{F\}, \{G\} \},$$

$$T^C = \{ \{A, B, C\}^C, \{D, E\}^C, \{F, G\}^C, \{A\}^C, \{B\}^C, \{C\}^C, \{D\}^C, \{E\}^C, \{F\}^C, \{G\}^C \},$$

$$\begin{aligned} \theta &= \{A, B, C, D, E, F, G\}, \\ m_T &= m_{ABC} \oplus m_{DE} \oplus m_{FG} \oplus m_A \oplus m_B \oplus m_C \oplus m_D \\ &\quad \oplus m_E \oplus m_F \oplus m_G, \\ m_T^C &= m_{ABC}^C \oplus m_{DE}^C \oplus m_{FG}^C \oplus m_A^C \oplus m_B^C \oplus \\ &\quad m_C^C \oplus m_D^C \oplus m_E^C \oplus m_F^C \oplus m_G^C \text{ 이다.} \end{aligned}$$

[Lemma 1] 징후 결합시 T만을 고려하는 경우보다 T 및 T^C를 같이 고려하는 경우 근사 오차를 줄일 수 있다.

(증명) 각 가설에 대한 불확증 징후 결합은 하나씩 차례로 결합되므로 임의의 가설 A에 대한 불확증 징후의 결합시 어느 가설 X의 bpa는 T를 고려하는 경우 $m_T \oplus m_A^C(X)$, TUT^C를 고려하는 경우 $m_{TUT^C} \oplus m_A^C(X)$ 로 나타낼 수 있다. 이때 X와 A 사이의 관계에 따라 $X \subseteq A$, $X \cap A = \emptyset$, $X \supset A$ 등 3가지로 나누어 생각할 수 있다. 각 경우에 대해 TUT^C를 고려하는 경우가 T만을 고려하는 경우보다 근사 오차가 같거나 같음을 보이겠다

(1) $X \subseteq A$ 인 경우

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T m_T(X) m_A^C(\theta), \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C m_T(X) m_A^C(\theta) \end{aligned}$$

이다. 이때에는 값의 상향 배당이 없으므로 두 경우 모두 오차가 발생하지 않는다

주어진 예제에서, T만을 고려하는 경우에 이와 같은 형태로 징후가 결합하는 예는 다음과 같다(A = {DE}, X = {D}라 하자).

$m_{DE}^C \setminus m_T$	{ABC}{DE}{FG}{A}{B}{C}{D}{E}{F}{G}{\theta}
{DE} ^C	{ABC} \emptyset {FG}{A}{B}{C} \emptyset \emptyset {F}{G}{DE} ^C
{\theta}	{ABC}{DE}{FG}{A}{B}{C}{D}{E}{F}{G}{\theta}

즉, $m_{DE}^C(\{\theta\})$ 와 $m_T(\{D\})$ 가 구해진다

(2) $X \cap A = \emptyset$ 인 경우

(i) $X \cup A \in T$ 인 경우

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T [m_T(X) + m_T(X \cup A) m_A^C(A^C)], \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C [m_{TUT^C}(X) + m_{TUT^C}(X \cup A) m_A^C(A^C)], \end{aligned}$$

이므로 두 경우 근사 오차가 없다.

(ii) $X \cup A \notin T$, $X \cup A \in TUT^C$ 인 경우

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T m_T(X), \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C [m_{TUT^C}(X) + m_{TUT^C}(X \cup A) m_A^C(A^C)], \end{aligned}$$

이므로 TUT^C를 고려하는 경우 $K_T^C [m_{TUT^C}(X \cup A) m_A^C(A^C)]$ 만큼의 오차를 줄일 수 있다.

(iii) $X \cup A \notin T$, $X \cup A \notin TUT^C$ 인 경우

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T m_T(X), \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C m_{TUT^C}(X) \end{aligned}$$

이므로 두 경우 근사 오차가 없다.

(3) $X \supset A$ 인 경우

(i) $X \cap A^C \in T$ 인 경우

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T m_T(X), \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C m_{TUT^C}(X) m_A^C(\theta) \end{aligned}$$

이므로 두 경우 근사 오차가 없다.

(ii) $X \cap A^C \notin T$, $X \cap A^C \in TUT^C$ 인 경우

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T m_T(X), \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C m_{TUT^C}(X) m_A^C(\theta) \end{aligned}$$

이므로 두 경우 근사 오차가 없다

(iii) $X \cap A^C \notin T$, $X \cap A^C \notin TUT^C$ 인 경우 Y가 임의의 가설을 나타낼 때,

$$\begin{aligned} m_T \oplus m_A^C(X) &= K_T m_T(X), \dots \dots \dots (8) \\ m_{TUT^C} \oplus m_A^C(X) &= K_T^C [m_{TUT^C}(X) m_A^C(\theta) + \\ &\quad m_{TUT^C}(Y) m_A^C(A^C) / N] \\ &\quad \dots \dots \dots (9) \end{aligned}$$

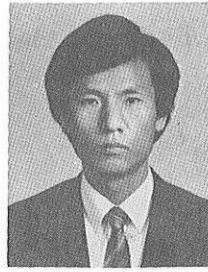
이므로 TUT^C를 고려하는 경우에 (식 8) - (식 9)의 절댓치 만큼의 근사 오차를 줄일 수 있다

위의 (1), (2), (3)에 의해 TUT^C를 고려하는 경우 T만을 고려하는 경우보다 근사 오차를 줄일 수 있다.

참 고 문 헌

1. Duda, R., Hart, P., and Nilsson, N., Subjective Bayesian Methods for Rule-Based Infer-

- ence Systems, in: *Proceedings 1976 National Computer Conference*, (AFIPS, 1976) 1075-1082.
2. Shortliffe, E. H. and Buchanan, B. G., A Model of Inexact Reasoning in Medicine, in B. G. Buchanan and E. H. Shortliffe, (Eds.), *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* (Addison-Wesley, Reading, MA, 1984) 233-262.
 3. Adams, J. B., Probabilistic Reasoning and Certainty Factors, in B. G. Buchanan and E. H. Shortliffe, (Eds.), *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* (Addison-Wesley, Reading, MA, 1984) 263-271.
 4. Shafer, G., *A Mathematical Theory of Evidence* (Princeton University Press, Princeton, NJ, 1976).
 5. Gordon, J. and Shortliffe, E. H., The Dempster-Shafer Theory of Evidence, in B. G. Buchanan and E. H. Shortliffe, (Eds.), *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* (Addison-Wesley, Reading, MA, 1984) 272-292.
 6. Barnett, J. A., Computational Methods for a Mathematical Theory of Evidence, in *Proceedings 7th International Joint Conference on Artificial Intelligence*, Vancouver, (BC 1981) 868-875.
 7. Gordon, J. and Shortliffe, E. H., A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space, *Artificial Intelligence* 26 (1985) 323-357.



이 종 만

- 1982년 육군사관학교 전자공학과 졸업
- 1987년 한국과학기술원 전산학과에서 석사학위 취득
- 1987년 한국과학기술원 전산학과 박사과정 입학, 재학중

주 관심분야 : 전문가 시스템에서의 불확실성의 처리, 복합 교장 진단등



김진형

- 1971년 서울대학교 공과대학 졸업
- 1979년 UCLA 전산학 석사학위 취득
- 1983년 UCLA 전산학 박사학위 취득
- 1973년~1976년 KIST 전

산실 연구원

- 1976년~1977년 미국 California 도로국 연구원
 - 1981년~1985년 미국 Hughes 인공지능 센터 선임 연구원
 - 1985년~현재 한국과학기술원 전산학과 조교수
 - 1986년~1988년 본 학회 산하 인공지능 연구회 위원장
- 주요 관심분야는 컴퓨터 시각 (Computer Vision), 전문가 시스템 (Expert System), 저능형 Man-Machine Interface 등임.