Ch03_ 규칙기반 전문가 시스템에서 불확실성 다루기

이 장에서 다룰 내용

- ❖ 01_불확실성이란?
- ❖ 02_기본 확률 이론
- ❖ 03_베이즈 추론
- ❖ 04_FORECAST: 베이즈 증거 축적
- ❖ 05_베이즈 방법의 치우침(bias)
- ❖ 06_확신도 이론과 증거 추론
- ❖ 07_FORECAST: 확신도의 응용 사례
- ❖ 08_베이즈 추론과 확신도의 비교
- ❖ 09_요약

01_불확실성이란?

❖ 불확실성이란?

- 전문가(사람)가 사용하는 정보의 공통적인 특징 중 하나는 불완전성임.
- 정보란 '불완전'할 수 있고, '모순'되기도 하며, '불확실'할 수도 있음.
- 어떤 정보는 경우에 따라 이 세가지 특성을 모두 띄기도 함.
- 주어진 정보가 문제를 푸는 데 적합하지 않을 수도 있음.
- 전문가는 이런 단점을 극복하고 올바른 판단과 결정을 내릴 수 있음.
- 전문가 시스템도 불확실성을 잘 다루어 타당한 결론을 도출할 수 있어야 함.

01_불확실성이란?

- 전문가 시스템에서 말하는 불확실성이란
  - 불확실성(uncertainty)이란 "확실하고 믿을 만한 결론에 도달하기 위한 정확한 정보의 부족"으로 정의할 수 있음.
  - 고전 논리학에서는 정확한 추론만 허용되는데, 이 논리에서 정확한 지식은 늘 존재하여 항상 흑백논리가 적용된다고 가정함.
  - 고전 논리학에서의 흑백논리 예:

IF	A는 참이다
THEN	A는 거짓이 아니다

IF	B는 거짓이다
THEN	B는 참이 아니다

  - 전문가 시스템이 다루는 대부분의 실세계 문제는 명확하게 분리된 지식을 제공하지 않음.
  - 우리가 사용하는 정보에는 종종 부정확하거나 불완전하기도 하고, 측정할 수 없는 데이터도 포함되어 있음.

01_불확실성이란?

❖ 전문가 시스템에서 지식이 불확실해지는 이유

- 상관관계가 취약한 함축
  - 규칙기반 전문가 시스템은 상관관계가 약한 함축과 애매한 결합으로 인해 종종 고생함.
  - 주제 전문가와 지식 공학자는 규칙에서 IF(조건)와 THEN(취해야 할 행동) 사이의 구체적인 상관관계를 만드는, 힘들고 희망이 없어 보이는 작업을 수행.
  - 따라서 전문가 시스템은 상관관계를 수치화된 확신도로 정의하는 것과 같은 방법으로 모호한 관계를 다루는 능력을 갖추어야 함.

- 부정확한 언어
  - 자연어(natural language)는 본질적으로 모호하고 부정확함.
  - 일반 지식들은 종종(often), 때때로(sometimes), 자주(frequently), 거의 하지 않는(hardly ever) 등과 같은 단어로 빈도를 표현.
  - 지식을 생성 규칙의 IF-THEN 형식으로 표현하기 어려울 수 있음. 하지만 기술된 지식의 의미를 제대로 정량화한다면 전문가 시스템에서 사용할 수 있음.
  - 1944년에 레이 심슨은 고등학생 355명과 대학생들에게 빈도와 관련된 영어단어 20개를 의미의 정도에 따라 1~100 사이에 표기해 달라고 함(Simpson, 1944).
  - 1968년에 밀턴 하켈은 이 실험을 다시 수행(Hakel, 1968).

01_불확실성이란?

- 모호하고 부정확한 영단어들을 시간당 빈도수 단위로 정량화한 결과

레이 심슨 (1944)		밀턴 하켈 (1968)	
단어	평균값	단어	평균값
Always	99	Always	100
Very often	88	Very often	87
Usually	85	Usually	79
Often	78	Often	74
Generally	78	Rather often	74
Frequently	73	Frequently	72
Rather often	65	Generally	72
About as often as not	50	About as often as not	50
Now and then	20	Now and then	34
Sometimes	20	Sometimes	29
Occasionally	20	Occasionally	28
Once in a while	15	Once in a while	22
Not often	13	Not often	16
Usually not	10	Usually not	16
Seldom	10	Seldom	9
Hardly ever	7	Hardly ever	8
Very seldom	6	Very seldom	7
Rarely	5	Rarely	5
Almost never	3	Almost never	2
Never	0	Never	0

01_불확실성이란?

❖ 전문가 시스템에서 지식이 불확실해지는 이유

- 알려지지 않은 데이터
  - 데이터가 불완전하거나 없을 때 유일한 해법은 그 값을 '알려지지 않은' 것으로 받아들이고, 근사적인 추론을 진행.
- 여러 전문가의 관점 통합
  - 일반적으로 대규모 전문가 시스템에는 수많은 전문가의 지식과 경험이 통합되어 있음.
  - 전문가들은 좀처럼 동일한 결론을 얻지 못하며 대개 모순된 견해와 충돌하는 규칙을 만들어냄
  - 지식 공학자는 이 문제를 해결하기 위해 각 전문가에게 가중치를 준 후 통합된 결론을 계산해야 함.
  - 그러나 주제 전문가들도 해당 주제에 대해 경험한 수준이 서로 다르며, 가중치를 얻는 체계적인 방법도 전혀 없음.
- 모든 실세계 영역에는 부정확한 지식이 있으며, 모순되거나 불완전하기도 하고, 심지어 분실된 데이터가 있을 수 있음.
- 그러므로 전문가 시스템은 불확실성을 다룰 수 있어야 함
  - 규칙기반 전문가 시스템에서는 불확실성을 다루기 위해 여러 수치적/비수치적 방법을 개발해 옴.

❖ 확률의 정의

- 사건의 확률은 사건이 발생할 경우의 비율이며, 확률을 가능성에 대한 과학적인 척도로 정의할 수 있음.
- 수학적으로 0(완전히 불가능함)과 1(완전히 확실함) 사이의 범위에 있는 수치적인 지표로 확률을 표현할 수 있음.
- 성공과 실패의 확률은 다음과 같이 정의함.

$$p(\text{성공}) = \frac{\text{성공 횟수}}{\text{모든 경우의 수}} \tag{3.1}$$

$$p(\text{실패}) = \frac{\text{실패 횟수}}{\text{모든 경우의 수}} \tag{3.2}$$

- 만약, s가 성공할 경우의 수고, f가 실패할 경우의 수라면

$$p(\text{성공}) = \ p = \frac{s}{s+f} \tag{3.3}$$

$$p(\text{실패}) = q = \frac{f}{s+f} \tag{3.4}$$

$$p + q = 1 \tag{3.5}$$

❖ 확률의 정의

- 동전을 던지면 앞면이 나올 확률(s)은 뒷면이 나올 확률(f)과 같음.
  - 한 번 던지는 경우, $s = f = 1$
  - 따라서 앞면(뒷면)이 나올 확률은 0.5

- 주사위를 한 번 던져서6이 나올 확률을 구해보자.
  - 숫자6을 유일한 성공이라고 가정하면 주사위를 한 번 던져서 6이 나오지 않는 다섯 가지의 경우가 있음.
  - 6이 나올 확률

$$p = \frac{1}{1+5} = 0.1666$$

  - 6이 나오지 않을 확률

$$q = \frac{5}{1+5} = 0.8333$$

❖ 확률의 정의

- 독립 사건과 배타적 사건
  - 동시에 일어날 수 없는 사건.
  - 주사위 실험에서6과1이 나오는 두 사건은 상호 배타적, 한 번 던졌을 때6과1이 동시에 나올 수는 없기 때문

- 독립된 사건이 아니면 다른 사건이 발생할 가능성에 영향을 줄 수도 있음.
  - e.g., 주사위를 한 번 던지는데, 이번에는 1이 나오지 않는다는 사실을 알고 있을 때 6이 나올 확률을 생각해보자.
  - 여전히6이 나오지 않는 다섯 가지 경우가 있지만 이미1이 나오지 않는다는 사실을 알고 있으므로 이 중 하나를 제외할 수 있음.
  - 따라서 다음 식을 얻을 수 있음.

$$p = \frac{1}{1 + (5 - 1)}$$

❖ 확률의 정의

- 조건부 확률(conditional probability): 사건 B가 발생했을 때 사건 A가 발생할 확률.
  - A를 하나의 사건(event), B를 또 다른 하나의 사건이라고 하자.
  - 사건 A와 B가 상호 배타적이지 않고, 조건부로 다른 사건의 발생에 영향을 준다고 가정하자.
  - 사건 B가 발생했을 때 사건 A가 발생할 확률을 조건부 확률이라 함.

- 조건부 확률은 수학적으로 p(A|B)라 적을 수 있음.
  - 수직 바(|)는 사건이 주어졌음을 표현.
  - 확률을 '사건 B가 발생했을 때 사건A가 발생할 조건부 확률'로 해석할 수 있음.

$$p(A|B) = \frac{A\text{와 } B\text{가 발생할 경우의 수}}{B\text{가 발생할 경우의 수}} \tag{3.6}$$

❖ 조건부 확률

- A와 B가 발생할 수 있는 경우의 비율, 즉 A와 B가 둘 다 발생할 확률을 A와 B의 결합 확률(joint probability)이라고 함.
- 수학적으로는 p(A ∩ B)로 표현. B가 발생하는 비율은 B의 확률인 p(B). 따라서 다음 식이 성립

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \tag{3.7}$$

- 유사하게, 사건 A가 발생했을 때 사건B가 발생할 조건부 확률은 다음과 같음

$$p(B|A) = \frac{p(B \cap A)}{p(A)} \tag{3.8}$$

- 따라서 다음 식이 성립

$$p(B \cap A) = p(B|A) \times p(A) \tag{3.9}$$

❖ 베이즈 규칙

- 결합 확률은 교환 법칙이 성립하므로

$$p(A \cap B) = p(B \cap A)$$

- 그렇기 때문에

$$p(A \cap B) = p(B|A) \times p(A) \tag{3.10}$$

- 따라서 다음 식이 성립

$$p(A|B) = \frac{p(B|A) \times p(A)}{p(B)} \tag{3.11}$$

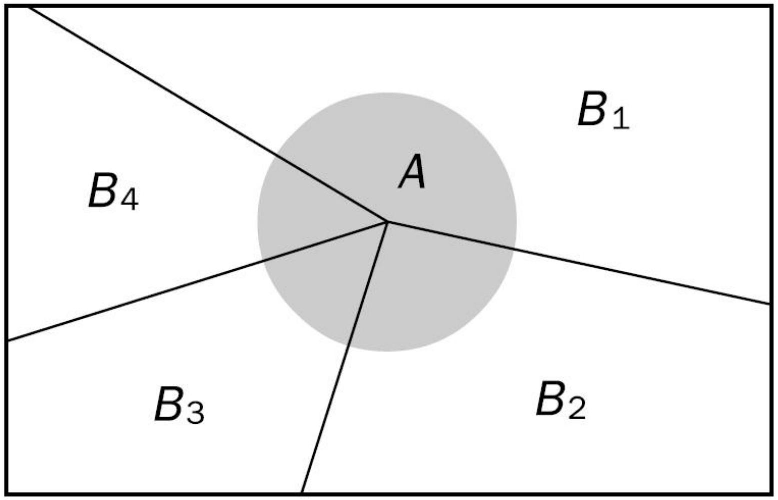
  - p(A|B)는 사건B가 발생했을 때 사건A가 발생하는 조건부 확률
  - p(B|A)는 사건A가 발생했을 때 사건 B가 발생할 조건부 확률
  - p(A)는 사건A가 발생할 확률, p(B)는 사건 B가 발생할 확률

- (3.11)은 **베이즈 규칙**(Bayes rule)으로 불림

❖ 결합 확률

- 지금까지 소개한 조건부 확률의 개념은 사건 A가 사건 B에 종속되어 있다는 점을 고려했음.
- 이 원리를 여러 상호 배타적인 사건들B1, B2, ... Bn에 종속적인 사건A로 확장할 수 있으므로 (3.7)에서 다음 식을 유도할 수 있음.

$$\sum_{i=1}^{n} p(A \cap B_i) = \sum_{i=1}^{n} p(A|B_i) \times p(B_i) \qquad (3.12)$$



❖ 결합 확률

- B_i가 빠짐없이 모두 망라되어 있고, (3.12)가 합쳐진다면 다음 식을 얻을 수 있음.

$$\sum_{i=1}^{n} p(A \cap B_i) = p(A) \tag{3.13}$$

  - (3.12)를 다음과 같이 조건부 확률식으로 바꿀 수 있음.

$$p(A) = \sum_{i=1}^{n} p(A|B_i) \times p(B_i) \tag{3.14}$$

  - 사건 A가 오직 상호 배타적인 두 사건에만 종속적으로 발생한다면

$$p(A) = p(A|B) \times p(B) + p(A|\neg B) \times p(\neg B) \tag{3.15}$$

  - 여기에서 ¬는 논리함수 NOT을 뜻함. 유사하게 다음 식을 얻을 수 있음.

$$p(B) = p(B|A) \times p(A) + p(B|\neg A) \times p(\neg A) \tag{3.16}$$

  - 이제 (3.16)을 베이즈 규칙 (3.11)에 대입하면 다음 식을 얻음.

$$p(A|B) = \frac{p(B|A) \times p(A)}{p(B|A) \times p(A) + p(B|\neg A) \times p(\neg A)} \tag{3.17}$$

- (3.17)은 전문가 시스템에서 불확실성을 다루기 위해 적용하는 확률 이론의 기초가 됨

확률 문제 #1

1. 조커를 뺀 트럼프 카드 52장중에서 카드 1장을 뽑았을 때, 다이아일 확률은 얼마인가?

2. 조커를 뺀 트럼프 카드 52장중에서 카드 2장을 뽑았을 때, 둘 다 다이아일 확률은 얼마인가?

3. 조커를 뺀 트럼프 카드 52장중에서 카드 1장을 뽑은 뒤, 어떤 카드인지 확인하지 않고 상자에 넣었다.

그리고 남은 카드를 잘 섞은 다음 3장을 뽑았는데, 3장 다 다이아였다.

이 때, 상자 안의 카드가 다이아일 확률은 얼마인가?

확률 문제 #2

집 근처에 새로 한 가족이 이사를 왔다. 아이가 2명 있다는 것을 알고 있지만, 아들인지 딸인지는 모른다.

1. 이웃집 부인에게 '딸이 있습니까'라고 물었더니 대답은 '네'였다. 다른 한 아이도 딸일 확률은 얼마인가?

2. 이웃집 부인에게 '큰 아이가 딸입니까?'라고 물었더니 대답은 '네'였다. 또 한명도 딸일 확률은 얼마인가?

3. 이웃집 부인이 딸을 1명 데리고 걷고 있는 것을 보았다. 다른 한 명의 아이도 딸일 확률은 얼마인가?

확률 문제 #3

세 개의 문 중에 하나를 선택하여 문 뒤에 있는 선물을 가질 수 있는 게임쇼에 참가했다.

한 문 뒤에는 자동차가 있고, 나머지 두 문 뒤에는 염소가 있다.

이때 어떤 사람이 예를 들어 1번 문을 선택했을 때, 게임쇼 진행자는 3번 문을 열어 문뒤에 염소가 있음을 보여주면서 1번 대신 2번을 선택하겠냐고 물었다.

이때 원래 선택했던 번호를 바꾸는 것이 유리할까?

❖ 베이즈 추론

- 기반 지식에 있는 모든 규칙을 다음과 같은 형태로 표현한다고 가정하자.

IF E가 참이다
THEN H는 {확률 p로} 참이다

- 이 규칙은 사건E가 발생하고 사건 H가 발생할 확률이 p임을 뜻함.
- 단순히A와B대신H와 E를 사용함,
- 전문가 시스템에서 H는 보통 가설(hypothesis)을 나타내고, E는 가설을 지지하는 증거(evidence)를 나타냄.

사건 E가 발생했지만 사건 H가 발생했는지 알지 못한다면 어떨까
사건 H가 발생할 확률도 계산할 수 있을까

- (3.17)를 이용하여 이 확률 값을 계산할 수 있음.
- A와 B대신 H와 E를 사용하여 확률 값을 계산.

$$p(H|E) = \frac{p(E|H) \times p(H)}{p(E|H) \times p(H) + p(E|\neg H) \times p(\neg H)} \quad (3.18)$$

❖ 베이즈 추론

$$p(H|E) = \frac{p(E|H) \times p(H)}{p(E|H) \times p(H) + p(E|\neg H) \times p(\neg H)} \tag{3.18}$$

- p(H): 가설H가 참일 사전 확률
- p(E|H): 가설H가 참일 때 증거E를 이끌어낼 확률
- p(¬H): 가설H가 거짓일 사전 확률
- p(E|¬H): 가설H가 거짓인데도 증거E를 얻는 확률

- (3.18)은 가설 H의 확률 p(H)를 어떤 증거가 관찰되기 전에 정의해야 함을 나타냄.
- 전문가 시스템에서 문제를 풀려면 전문가들이 필요한 확률 값을 제공해야 함.
- 전문가는 가능한 가설 p(H)와 p(¬H)에 대한 사전 확률을 결정하고, 가설 H가 참일 때 증거 E를 관찰할 조건부 확률 p(E|H)와 가설 H가 거짓인 경우에 대한 조건부 확률 p(E|¬H)를 결정.
- 사용자는 관찰한 증거에 대한 정보를 제공하고, 전문가 시스템은 사용자가 제공한 증거 E의 관점에서 가설 H에 대한 확률 p(H|E)를 계산.
- 확률 p(H|E)는 증거 E를 관찰했을 때 가설 H의 **사후 확률**이라 함.

단일 증거 E에 기반을 둔 전문가가 단일 가설이 아닌 가설 $H_1, H_2, \ldots, H_m$을 선택한다면 어떨까
또는 증거 $E_1, E_2, \ldots, E_n$ 이 주어지면 전문가도 여러 가설을 만들 수 있을까

- **가설** $H_1, H_2, \ldots, H_m$ **과 증거** $E_1, E_2, \ldots, E_n$ **을 이용하여** (3.18)**을 일반화할 수 있음**.
- **그러나 증거와 가설 모두 상호 배타적이고 총망라된 것이어야 함**.
- **단일 증거** E**와 가설** $H_1, H_2, \ldots, H_m$ **에 관한 식**:

$$p(H_i|E) = \frac{p(E|H_i) \times p(H_i)}{\sum_{k=1}^{m} p(E|H_k) \times p(H_k)} \tag{3.19}$$

- **증거** $E_1, E_2, \ldots, E_n$ **과 가설** $H_1, H_2, \ldots, H_m$ **에 관한 식**:

$$p(H_i|E_1E_2\ldots E_n) = \frac{p(E_1E_2\ldots E_n|H_i) \times p(H_i)}{\sum_{k=1}^{m} p(E_1E_2\ldots E_n|H_k) \times p(H_k)} \tag{3.20}$$

03_베이즈 추론

- (3.20)을 적용하려면, 모든 가설에 대해 모든 증거 조합의 조건부 확률을 알아야 함.
- 이 요건은 전문가에게 엄청난 부담만 주고, 실제로 작업하기도 어려움.
- 그러므로 전문가 시스템에서는 증거가 지나치게 세밀해지는 것을 피하고, 다른 증거 간의 조건부 독립성을 띠게 해야 함.
- 실제적으로는 값을 구할 수 없는 (3.20) 대신 다음 식을 얻음.

$$p(H_i|E_1E_2\ldots E_n) = \frac{p(E_1|H_i) \times p(E_2|H_i) \times \ldots \times p(E_n|H_i) \times p(H_i)}{\sum_{k=1}^{m} p(E_1|H_k) \times p(E_2|H_k) \times \ldots \times p(E_n|H_k) \times p(H_k)} \tag{3.21}$$

전문가 시스템은 어떻게 모든 사후 확률을 계산하고, 잠정적으로 참인 가설에 순위를 매길 수 있을까

- 세 개의 조건부 독립인 증거 E_1, E_2, E_3이 주어지면 전문가는 세 개의 상호 배타적이고 총 망라된 가설 H_1, H_2, H_3 을 만들고, 가설에 대한 사전확률 $p(H_1)$, $p(H_2)$, $p(H_3)$을 제공한다고 가정하자.
- 또한 가설이 주어졌을 때, 전문가는 가능한 모든 가설에 대해 증거를 관찰할 조건부 확률을 제공함,
- 아래 표는 전문가가 제공한 사전 확률과 조건부 확률을 보여줌.

확률	가설		
	$i = 1$	$i = 2$	$i = 3$
$P(H_i)$	0.40	0.35	0.25
$P(E_1\|H_i)$	0.3	0.8	0.5
$P(E_2\|H_i)$	0.9	0.0	0.7
$P(E_3\|H_i)$	0.6	0.7	0.9

- 먼저 증거 E_3 을 관찰했다고 가정하자.
- 전문가 시스템은 (3.19)에 따라 모든 가설에 대해 사후 확률을 계산함.

$$p(H_i|E_3) = \frac{p(E_3|H_i) \times p(H_i)}{\sum_{k=1}^{3} p(E_3|H_k) \times p(H_k)} \qquad i = 1, 2, 3$$

- 따라서 다음 값을 얻음.

$$p(H_1|E_3) = \frac{0.6 \times 0.40}{0.6 \times 0.40 + 0.7 \times 0.35 + 0.9 \times 0.25} = 0.34$$

$$p(H_2|E_3) = \frac{0.7 \times 0.35}{0.6 \times 0.40 + 0.7 \times 0.35 + 0.9 \times 0.25} = 0.34$$

$$p(H_3|E_3) = \frac{0.9 \times 0.25}{0.6 \times 0.40 + 0.7 \times 0.35 + 0.9 \times 0.25} = 0.32$$

- 증거 E_3을 관찰한 후에 가설 H_1에 대한 신뢰는 감소하여 가설 H_2 에 대한신뢰와 같아짐.
- 반면 가설 H_3에 대한 신뢰는 증가하여 가설 H_1 과 H_2 에 대한 신뢰에 거의 근접함.

03_베이즈 추론

- 증거 E_1을 관찰했다고 가정하자. (3.21)로 계산한 사후 확률은 다음과 같음

$$p(H_i|E_1E_3) = \frac{p(E_1|H_i) \times p(E_3|H_i) \times p(H_i)}{\sum_{k=1}^{3} p(E_1|H_k) \times p(E_3|H_k) \times p(H_k)} \qquad i = 1, 2, 3$$

- 따라서 다음 값을 얻음

$$p(H_1|E_1E_3) = \frac{0.3 \times 0.6 \times 0.40}{0.3 \times 0.6 \times 0.40 + 0.8 \times 0.7 \times 0.35 + 0.5 \times 0.9 \times 0.25} = 0.19$$

$$p(H_2|E_1E_3) = \frac{0.8 \times 0.7 \times 0.35}{0.3 \times 0.6 \times 0.40 + 0.8 \times 0.7 \times 0.35 + 0.5 \times 0.9 \times 0.25} = 0.52$$

$$p(H_3|E_1E_3) = \frac{0.5 \times 0.9 \times 0.25}{0.3 \times 0.6 \times 0.40 + 0.8 \times 0.7 \times 0.35 + 0.5 \times 0.9 \times 0.25} = 0.29$$

  - 이제 가설 H_2가 가장 믿을만 해 보임.
  - 반면 가설 H_1 에 대한 신뢰는 현저히 감소. 마찬가지로 증거 E_2 를 관찰한 후, 전문가 시스템은 모든 가설에 대해 사후 확률을 계산.

$$p(H_i|E_1E_2E_3) = \frac{p(E_1|H_i) \times p(E_2|H_i) \times p(E_3|H_i) \times p(H_i)}{\sum_{k=1}^{3} p(E_1|H_k) \times p(E_2|H_k) \times p(E_3|H_k) \times p(H_k)} \qquad i = 1, 2, 3$$

- 그러면 다음과 같은 값을 얻음.

$$p(H_1|E_1E_2E_3)=\frac{0.3\times0.9\times0.6\times0.40}{0.3\times0.9\times0.6\times0.40+0.8\times0.0\times0.7\times0.35+0.5\times0.7\times0.9\times0.25}$$
$$=0.45$$

$$p(H_2|E_1E_2E_3)=\frac{0.8\times0.0\times0.7\times0.35}{0.3\times0.9\times0.6\times0.40+0.8\times0.0\times0.7\times0.35+0.5\times0.7\times0.9\times0.25}$$
$$=0$$

$$p(H_3|E_1E_2E_3)=\frac{0.5\times0.7\times0.9\times0.25}{0.3\times0.9\times0.6\times0.40+0.8\times0.0\times0.7\times0.35+0.5\times0.7\times0.9\times0.25}$$
$$=0.55$$

- 전문가가 제시한 최초 순위가 H_1, H_2, H_3 이었다 할지라도 모든 증거(E_1, E_2, E_3)를 관찰한 후에는 가설 H_1 과 H_3 만 고려 대상으로 남음.
- 이제 가설 H_2 는 고려하지 않아도 된다. 이때 가설 H_3 가 가설 H_1 보다 더 믿을 만해 보인다는 점에 주목 해야함.
- 광물 탐사 전문가 시스템인 PROSPECTOR는 확률 p(H|E)를 계산하기 위해 베이즈 규칙을 사용.
  - 시스템 전체에 걸쳐 불확실한 지식에서 추론을 수행한 첫 번째 시스템

❖ 일기예보 전문가 시스템

- 일기예보 전문가 시스템은 내일 비가 올지 안 올지를 알아내야 하므로 기상청에서 실제 데이터를 얻어와야 함.
- [표 3-3]은 1982년 3월의 런던 날씨를 요약한 것. 강우량이 0이면 맑은 날씨.
- 전문가 시스템은 '내일은 비가 온다'와 '내일은 맑다'라는 두 개의 결과를 내놓고, 그 확률을 제공해야 함,
- 베이즈 규칙을 적용하기 위해서는 가정들에 대한 사전 확률을 제시해야 함.

규칙 1

IF 오늘은 비가 온다

THEN 내일은 비가 온다

규칙 2

IF 오늘은 맑다

THEN 내일은 맑다

- 이 규칙을 사용하면 비오는 날씨가 맑은 날씨 바로 앞에 있거나 맑은 날씨가 비오는 날씨 앞에 있을 때마다 실수를 할 것임.

[표 3-3] 1982년 3월 런던 날씨에 대한 요약 정보

날짜	최저기온(℃)	최고기온(℃)	강우량(mm)	일조시간	실제 날씨	기상예보
1	9.4	11.0	17.5	3.2	비	-
2	4.2	12.5	4.1	6.2	비	비
3	7.6	11.2	7.7	1.1	비	비
4	5.7	10.5	0.0	4.3	맑음	비*
5	3.0	12.0	0.0	9.5	맑음	맑음
6	4.4	9.6	0.0	3.5	맑음	맑음
7	4.8	9.4	4.6	10.1	비	비
8	1.8	9.2	5.5	7.8	비	비
9	2.4	10.2	4.8	4.1	비	비
10	5.5	12.7	4.2	3.8	비	비
11	3.7	10.9	4.4	9.2	비	비
12	5.9	10.0	4.8	7.1	비	비
13	3.0	11.9	0.0	8.3	맑음	비*
14	5.4	12.1	4.8	1.8	비	맑음*
15	8.8	9.1	8.8	0.0	비	비

16	2.4	8.4	3.0	3.1	비	비
17	4.3	10.8	0.0	4.3	맑음	맑음
18	3.4	11.1	4.2	6.6	비	비
19	4.4	8.4	5.4	0.7	비	비
20	5.1	7.9	3.0	0.1	비	비
21	4.4	7.3	0.0	0.0	맑음	맑음
22	5.6	14.0	0.0	6.8	맑음	맑음
23	5.7	14.0	0.0	8.8	맑음	맑음
24	2.9	13.9	0.0	9.5	맑음	맑음
25	5.8	16.4	0.0	10.3	맑음	맑음
26	3.9	17.0	0.0	9.9	맑음	맑음
27	3.8	18.3	0.0	8.3	맑음	맑음
28	5.8	15.4	3.2	7.0	비	맑음*
29	6.7	8.8	0.0	4.2	맑음	맑음
30	4.5	9.6	4.8	8.8	비	비
31	4.6	9.6	3.2	4.2	비	비

* 기상예보가 틀렸음

❖ 일기예보 전문가 시스템

규칙 1

IF 오늘은 비가 온다 {LS 2.5 LN 0.6}

THEN 내일은 비가 온다 {사전 0.5}

규칙 2

IF 오늘은 맑다 {LS 1.6 LN 0.4}

THEN 내일은 맑다 {사전 0.5}

- LS(Likelihood of Sufficiency) 값은 증거 E가 있을 때 전문가가 가설 H를 신뢰하는 정도를 의미하는데, 이를 충분 가능성이라 함.
- 충분 가능성은 p(E|¬H)에 대한p(E|H)의 비율로 정의

$$LS = \frac{p(E|H)}{p(E|\neg H)} \tag{3.22}$$

- 위의 경우에 LS는 내일 비가 온다면 오늘 비가 올 확률을 내일 비가 오지 않는다면 오늘 비가 올 확률로 나눈 것

$$LS = \frac{p(\text{오늘은 비가 온다} \mid \text{내일은 비가 온다})}{p(\text{오늘은 비가 온다} \mid \text{내일은 맑다})}$$

04_FORECAST: 베이즈 증거 축적

- LN(Likelihood of Necessity) 값은 증거 E가 없을 때 가설 H에 대한 불신의 정도
- LN을 필요가능성이라 하고 다음과 같이 정의

$$LN = \frac{p(\neg E|H)}{p(\neg E|\neg H)} \qquad (3.23)$$

- 위 일기예보의 예에서 LN은 내일 비가 온다면 오늘 비가 오지 않을 확률을 내일 비가 오지 않으면 오늘 비가 오지 않을 확률로 나눈 것

$$LN = \frac{p(\text{오늘은 맑다} \mid \text{내일은 비가 온다})}{p(\text{오늘은 맑다} \mid \text{내일은 맑다})}$$

- LS로부터 LN을 유도할 수 없다는 사실을 기억해야 하며, 주제 전문가는 두 가지 값을 각각 제공해야 함.

04_FORECAST: 베이즈 증거 축적

주제 전문가는 충분 가능성과 필요 가능성의 값을 어떻게 결정하는가
전문가가 조건부 확률을 다루어야 하는가

- LS와 LN 값을 제공하기 위해 전문가가 정확한 조건부 확률 값을 결정할 필요는 없음. 전문가는 그 가능성을 결정하면 됨.
- LS 값이 크다(LS >> 1)라는 말은 증거가 관찰되면 그 규칙이 가설을 강력하게 뒷받침하며, LN 값이 작다(0 < LN < 1)는 말은 증거가 없다면 그 규칙이 가설을 강력하게 반대한다는 의미
- 조건부 확률은 LS와 LN 값으로 쉽게 계산할 수 있으므로, 증거를 전파할 때 베이즈 규칙을 사용할 수 있음.
  - 규칙 1은 오늘 비가 오고 있다면 내일 비가 올 확률이 매우 높다는 것을 뜻함(LS = 2.5).
  - 오늘 비가 오지 않는다 해도, 즉 맑다고 해도 내일 비가 올 확률이 약간은 있음(LN = 0.6).
  - 규칙2는 맑은 날을 다룸. 오늘 맑으면 내일 맑을 확률도 높음(LS = 1.6).
  - 오늘 비가 오고 있다면 내일 비가 오게 될 확률이 오늘 맑으면 내일 맑을 확률보다 높음. 그 이유는 LS와LN의 값을 대개 주제 전문가가 정하기 때문. 지금 다루는 날씨의 예에서는 그 값들을 기상청이 발표한 통계 정보로 확정할 수도 있음.
  - 또한, 규칙 2는 오늘 비가 올 때에도 내일 날씨가 맑을확률을 알려줌(LS = 0.4).

전문가 시스템은 내일 비가 오거나 맑다는 것에 대한 전반적인 확률을 어떻게 얻는가

- 규칙기반 전문가 시스템에서는 후건의 사전 확률 p(H)를 사전 가능성(prior odds)으로 변환할 수 있음.

$$O(H) = \frac{p(H)}{1 - p(H)} \tag{3.24}$$

- 사전 확률은 후건의 불확실성이 최초로 적용될 때만 쓰임.
- 그런 후에 차례로 사후 가능성을 구하는데, 규칙의 전건이 참일 때는 LS를 이용해서, 전건이 거짓일 때는 LN을 이용해서 사전 가능성을 갱신함.

$$O(H|E) = LS \times O(H) \tag{3.25}$$

$$O(H|\neg E) = LN \times O(H) \tag{3.26}$$

04_FORECAST: 베이즈 증거 축적

- 사후 가능성은 사후 확률을 구할 때 쓰임.

$$p(H|E) = \frac{O(H|E)}{1 + O(H|E)} \quad (3.27)$$

$$p(H|\neg E) = \frac{O(H|\neg E)}{1 + O(H|\neg E)} \quad (3.28)$$

- 런던 날씨 예는 이 방식이 어떻게 동작하는지 보여준다.
- 사용자가 오늘은 비가 온다고 알려주었다고 가정.
- 규칙 1이 점화되고, '내일은 비가 온다'의 사전 확률은 사전 가능성으로 변환됨.

$$O(\text{내일은 비가 온다}) = \frac{0.5}{1 - 0.5} = 1.0$$

- 오늘은 비가 온다는 증거는 그 가능성을 2.5만큼 증가시키고, 그로 인해 확률은 0.5에서 0.71로 높아짐.

$$O(\text{내일은 비가 온다} \mid \text{오늘은 비가 온다}) = 2.5 \times 1.0 = 2.5$$

$$p(\text{내일은 비가 온다} \mid \text{오늘은 비가 온다}) = \frac{2.5}{1 + 2.5} = 0.71$$

04_FORECAST: 베이즈 증거 축적

- 규칙 2 또한 점화된다. '내일은 맑다'의 사전 확률이 사전 가능성으로 바뀌지만, 오늘은 비가 온다는 증거가 그 가능성을 0.4만큼 낮춤.
- 이로 인해 '내일은 맑다'의 확률이 0.5에서 0.29로 줄어듦.

$$O(\text{내일은 맑다}) = \frac{0.5}{1-0.5} = 1.0$$

$$O(\text{내일은 맑다} \mid \text{오늘은 비가 온다}) = 0.4 \times 1.0 = 0.4$$

$$p(\text{내일은 맑다} \mid \text{오늘은 비가 온다}) = \frac{0.4}{1+0.4} = 0.29$$

- 따라서 오늘 비가 오는 중이면 내일 비가 올 확률이71%고 내일 맑을확률은29%.
- 또 사용자가 오늘은 맑다라고 입력했다고 가정해보자.
- 동일한 계산을 거치면 내일 맑을 확률은62%고 내일 비가 올 확률은38%가 됨.

❖ 날씨 분석을 위한 기반 지식

- [표 3-3] 참조

/* FORECAST: 베이즈 증거 축적

control 베이즈

규칙 1
IF 오늘은 비가 온다 {LS 2.5 LN 0.6}
THEN 내일은 비가 온다 {사전 확률 0.5}

규칙 2
IF 오늘은 맑다 {LS 1.6 LN 0.4}
THEN 내일은 맑다 {사전 확률 0.5}

규칙 3
IF 오늘은 비가 온다
AND 강우량이 적다 {LS 10 LN 1}
THEN 내일은 맑다 {사전 확률 0.5}

/* SEEK 지시어는 규칙 집합의 목표를 설정한다.

seek 내일

규칙 4
IF 오늘은 비가 온다
AND 강우량이 적다
AND 기온이 낮다 {LS 1.5 LN 1}
THEN 내일은 맑다 {사전 확률 0.5}

규칙 5
IF 오늘은 맑다
AND 기온이 높다 {LS 2 LN 0.9}
THEN 내일은 비가 온다 {사전 확률 0.5}

규칙 6
IF 오늘은 맑다
AND 기온이 높다
AND 하늘이 흐리다 {LS 5 LN 1}
THEN 내일은 비가 온다 {사전 확률 0.5}

❖ 전문가 시스템: 대화

- 사용자가 제공한 정보를 바탕으로 전문가 시스템은 내일 비가 올지를 결정.
- 강우량이 4.1mm보다 적으면 강우량이 적고, 하루 동안의 평균기온이 7.08℃ 이하면 기온이 낮고, 7.08℃ 보다 높으면 기온이 높다고 가정.
- 일조시간이 하루4.6시간보다 적으면 흐리다는 것을 의미.

(질문) ⇒ (답변)
오늘의 날씨가 어떻습니까? ⇒ 비가 온다

규칙 1

IF 오늘은 비가 온다 {LS 2.5 LN 0.6}

THEN 내일은 비가 온다 {사전 확률 0.5}

$$O(\text{내일은 비가 온다}) = \frac{0.5}{1 - 0.5} = 1.0$$

$$O(\text{내일은 비가 온다} \mid \text{오늘은 비가 온다}) = 2.5 \times 1.0 = 2.5$$

$$p(\text{내일은 비가 온다} \mid \text{오늘은 비가 온다}) = \frac{2.5}{1 + 2.5} = 0.71$$

내일은 비가 온다 {0.71}

규칙 2

IF 오늘은 맑다 {LS 1.6 LN 0.4}

THEN 내일은 맑다 {사전 확률 0.5}

$$O(\text{내일은 맑다}) = \frac{0.5}{1 - 0.5} = 1.0$$

$$O(\text{내일은 맑다} \mid \text{오늘은 비가 온다}) = 0.4 \times 1.0 = 0.4$$

$$p(\text{내일은 맑다} \mid \text{오늘은 비가 온다}) = \frac{0.4}{1 + 0.4} = 0.29$$

내일은 비가 온다 {0.71}

내일은 맑다 {0.29}

04_FORECAST: 베이즈 증거 축적

- 오늘의 강우량은 어떻습니까?
  ⇒ 적다

규칙 3

IF 오늘은 비가 온다
AND 강우량이 적다 {LS 10 LN 1}
THEN 내일은 맑다 {사전 확률 0.5}

$$O(\text{내일은 맑다}) = \frac{0.29}{1 - 0.29} = 0.41$$

$$O(\text{내일은 맑다} \mid \text{오늘은 비가 온다} \cap \text{강우량이 적다}) = 10 \times 0.41 = 4.1$$

$$p(\text{내일은 맑다} \mid \text{오늘은 비가 온다} \cap \text{강우량이 적다}) = \frac{4.1}{1 + 4.1} = 0.80$$

내일은 맑다 {0.80}
내일은 비가 온다 {0.71}

- 오늘은 기온은 어떻습니까?
  ⇒ 낮다

규칙 4

IF 오늘은 비가 온다
AND 강우량이 적다
AND 기온이 낮다 {LS 1.5 LN 1}
THEN 내일은 맑다 {사전 확률 0.5}

$$O(\text{내일은 맑다}) = \frac{0.80}{1 - 0.80} = 4$$

$$O(\text{내일은 맑다} \mid \text{오늘은 비가 온다} \cap \text{강우량이 적다} \cap \text{기온이 낮다}) = 1.50 \times 4 = 6$$

$$p(\text{내일은 맑다} \mid \text{오늘은 비가 온다} \cap \text{강우량이 적다} \cap \text{기온이 낮다}) = \frac{6}{1 + 6} = 0.86$$

내일은 맑다 {0.86}
내일은 비가 온다 {0.71}

규칙 5

IF 오늘은 맑다
AND 기온이 높다 {LS 2 LN 0.9}
THEN 내일은 비가 온다 {사전 확률 0.5}

$$O(\text{내일은 비가 온다}) = \frac{0.71}{1 - 0.71} = 2.45$$

$$O(\text{내일은 비가 온다} \mid \text{오늘은 맑지 않다} \cap \text{기온이 높지 않다}) = 0.9 \times 2.46 = 2.21$$

$$p(\text{내일은 비가 온다} \mid \text{오늘은 맑지 않다} \cap \text{기온이 높지 않다}) = \frac{2.21}{1 + 2.21} = 0.69$$

내일은 맑다 {0.86}
내일은 비가 온다 {0.69}

04_FORECAST: 베이즈 증거 축적

- **오늘 하늘은 어떻습니까? ⇒ 흐리다**

규칙 **6**

IF 오늘은 맑다

AND 기온이 높다

AND 하늘이 흐리다 {**LS 5 LN 1**}

THEN 내일은 비가 온다 {사전 확률 0.5}

$$O(\text{내일은 비가 온다}) = \frac{0.69}{1 - 0.69} = 2.23$$

$$O(\text{내일은 비가 온다} \mid \text{오늘은 맑지 않다} \cap \text{기온이 높지 않다} \cap \text{하늘이 흐리다}) = 1.0 \times 2.23 = 2.23$$

$$p(\text{내일은 비가 온다} \mid \text{오늘은 맑지 않다} \cap \text{기온이 높지 않다} \cap \text{하늘이 흐리다}) = \frac{2.23}{1 + 2.23} = 0.69$$

오늘은 맑다	{0.86}
오늘은 비가 온다	{0.69}

- **이는 일어날 가능성이 있는 두 개의 가설, 내일은 '맑다'와 내일은 '비가 온다' 중 전자의 가능성이 높다는 것을 의미한다.**

05_베이즈 방법의 치우침(bias)

❖ 베이즈 방법의 치우침(bias)

- 베이즈 추론의 주요 입력에는 확률 값이 들어가야 하고, 이 값을 할당하는 일은 보통 인간의 판단과 관계가 있음.
- 그러나 심리학 연구에 따르면 인간은 베이즈 규칙과 모순되지 않는 확률 값을 제대로 도출하지 못하고, 이는 조건 부 확률 값이 전문가가 제시한 사전 확률과 모순될 수도 있다는 것을 암시함.
- 시동을 걸었는데도 출발하지 않고 이상한 소리를 내는 차가 있다고 가정해보자.
- 차가 이상한 소리를 낼 경우, 시동 장치가 고장 났을 조건부 확률을 다음과 같이 표현할 수 있음.

IF 증상은 '이상한 소리'다
THEN 시동 장치가 고장 났다 {확률 0.7}

- 차가 이상한 소리를 내는데도 시동 장치가 고장 나지 않을 조건부 확률

$$p(\text{시동 장치가 고장 나지 않았다} \mid \text{이상한 소리})$$
$$= p(\text{시동 장치가 정상이다} \mid \text{이상한 소리}) = 1 - 0.7 = 0.3$$

- 따라서 다음의 대응 규칙을 얻는다.

IF 증상은 '이상한 소리'다
THEN 시동 장치가 정상이다 {확률 0.3}

❖ 베이즈 방법의 치우침(bias)

- 주제 전문가는 조건부 확률을 취급하지 않으며, 대개 숨겨진 내재 확률(hidden implicit probability) (여기서는0.3)을 고려하지 않음.
- 앞의 예에서, 유효한 통계 정보와 경험적 연구를 근거로 다음과 같은 규칙 두 개를 이끌어 낼 수 있음.

IF	시동 장치가 고장 났다
THEN	증상은 '이상한 소리'다 {확률 0.85}
IF	시동 장치가 고장 났다
THEN	증상은 '이상한 소리'가 아니다 {확률 0.15}

- 베이즈 규칙을 사용하려면 사전 확률, 즉 차가 출발하지 않을 때 시동 장치가 고장 났을 확률이 필요함.
- 전문가가 5%의 값을 제공했다고 가정해보자.
- 이제 베이즈 규칙을 적용하여 다음 식을 얻을 수 있음.

$$p(\text{시동 장치가 고장 났다} \mid \text{이상한 소리}) = \frac{0.85 \times 0.05}{0.85 \times 0.05 + 0.15 \times 0.95} = 0.23$$

- 계산한 수는 이 절의 처음에 주어진 전문가의 예측치 0.7보다 훨씬 낮음.

❖ 모순이 생기는 이유

- 모순의 이유는 전문가가 조건부 확률과 사전 확률을 정할 때 다른 가정을 했기 때문
- 사후 확률 p(시동 장치가 고장 났다 | 이상한 소리)에서부터 사전 확률 p(시동 장치가 고장 났다)까지 거꾸로 생각해보면 이를 조사할 수 있음.
  - 이 경우에, 다음과 같이 가정할 수 있음.

$$p(\text{시동 장치가 정상이다}) = 1 - p(\text{시동 장치가 고장 났다})$$

  - 베이즈 규칙으로부터 다음을 얻음.

$$p(H) = \frac{p(H|E) \times p(E|\neg H)}{p(H|E) \times p(E|\neg H) + p(E|H)[1 - p(H|E)]}$$

  - 여기서 각각의 확률은 다음과 같음.

$$p(H) = p(\text{시동 장치가 고장 났다})$$

$$p(H|E) = p(\text{시동 장치가 고장 났다} \mid \text{이상한 소리})$$

$$p(E|H) = p(\text{이상한 소리} \mid \text{시동 장치가 고장 났다})$$

$$p(E|\neg H) = p(\text{이상한 소리} \mid \text{시동 장치가 정상이다})$$

05_베이즈 방법의 치우침(bias)

❖ 모순이 생기는 이유

- 만일 p(시동 장치가 고장 났다 | 이상한 소리)의 확률을 전문가가 제공한 0.7로 고치면 사전 확률p(시동 장치가 고장 났다)의 값은 다음과 같이 될 것이다.

$$p(H) = \frac{0.7 \times 0.15}{0.7 \times 0.15 + 0.85 \times (1 - 0.7)} = 0.29$$

  - 이 값은 전문가가 제시한 5%보다 거의 6배 크다. 따라서 전문가는 사전 확률과 조건부 확률에 대해 매우 다른 추정을 한 것
  - 사실 전문가가 제공한 사전 확률 또한 충분 가능성(LS)이나 필요 가능성(LN)과 일치하지 않을 수있다. 따라서 이 문제를 처리하기 위한 몇 가지 방법이 제안됨.
  - 많이 사용하는 기법은 개별 선형 보간법 모델로, PROSPECTOR에 처음으로 적용

- 실제 베이즈 접근법을 사용하려면 어떤 가정과 부정 하에서 증거의 조건부 독립을 포함하여 많은 가정을 충족해야 함.
- 실세계 문제에서는 이런 가정을 제대로 만족하지 못하므로 몇몇 시스템만이 베이즈 추론에 기반을 두고 만들어졌음.
- 가장 많이 알려진 전문가 시스템은 광물 탐사를 수행하는 PROSPECTOR.

❖ 확신도 (cf, certainty factor) 이론과 증거 추론

- 확신도 이론은 베이즈 추론의 대안으로 가장 널리 알려짐.
- 이론의 기본 원리는 순환계 감염이나 뇌막염을 진단하고 치료하는 전문가 시스템인 MYCIN에 처음으로 도입.
  - MYCIN 개발자들은 의학 전문가들이 자신들의 신뢰도를 비논리적이고 수학적으로 일관성 없는 형태로 표현한다는 사실을 발견 했고, 문제 영역에 대한 믿을만한 통계 데이터도 없었음.
  - 이 때문에 MYCIN 팀은 고전적인 확률적 접근법을 사용하지 못하고, 확신도라는 전문가의 신뢰를 측정하는 수치를 도입하기로 결정했음.

- 확신도의 최대값은+1.0(확실히 참)이고, 최소값은-1.0(확실히 거짓)
- +값은 신뢰의 정도를 나타내고,-값은 불신의 정도를 나타냄.
- 확신도를 사용하는 전문가 시스템에서 기반지식은 다음의 문법을 포함한 규칙으로 이루어 짐.

IF ⟨증거⟩

THEN ⟨가설⟩ $\{cf\}$

  - *cf* 는 증거E가 발생했다고 할 때 가설H에 대한 신뢰를 나타냄.
  - 확신도 이론은 신뢰의 정도 MB(H,E)와 불신의 정도MD(H,E)를 나타내는 두 함수에 기초함.

❖ 확신도 이론과 증거 추론

- 신뢰와 불신의 정도는 사전 확률과 조건부 확률 형식으로 정의할 수 있음.

$$MB(H,E) = \begin{cases} 1 & \text{if } p(H) = 1 \\ \dfrac{max\,[p(H|E), p(H)] - p(H)}{max\,[1,0] - p(H)} & \text{otherwise} \end{cases} \quad (3.29)$$

$$MD(H,E) = \begin{cases} 1 & \text{if } p(H) = 0 \\ \dfrac{min\,[p(H|E), p(H)] - p(H)}{min\,[1,0] - p(H)} & \text{otherwise} \end{cases} \quad (3.30)$$

- 여기서 각각의 확률은 다음과 같다.

$p(H)$: 가설 H가 참일 사전 확률

$p(H \mid E)$: 증거 E가 주어졌을 때 가설 H가 참일 확률

  - MB(H, E)와 MD(H, E)에는 0~1 사이의 값이 들어감.
  - 가설 H에 대한 신뢰나 불신의 강도는 관찰된 증거 E의 종류에 따라 달라짐.
  - 어떤 사실은 신뢰의 강도를 높이지만, 어떤 것은 불신의 강도를 높임.

❖ 가설에 대한 신뢰 또는 불신의 강도 결정

- 하나의 수치, 확신도로 결합해서 설명하기 위해 식(3.31)을 사용

$$cf = \frac{MB(H,E) - MD(H,E)}{1 - min\,[MB(H,E), MD(H,E)]} \qquad (3.31)$$

  - *cf*는 가설 H에 대한 총체적 신뢰를 나타냄. MYCIN에서는 -1~+1 의 값을 갖게됨.

- MYCIN의 접근 방법을 예를 통해 살펴보자.

IF　　A는 X다
THEN　B는 Y다

  - 보통 전문가는 이 규칙이 성립하는지 100% 확신할 수 없음.
  - 또한 규칙의 IF 부분을 만족하여 객체A에X 값이 들어있는데, 객체B의 값이 Z인 경우를 관찰했다고 가정, 즉 통계에 반하는 불확실성이 나타난 것임.
  - 전문가는 보통A 값이X일 때 가능한 B 값 각각에 대해 확신도를 정함.
  - 따라서 위 규칙은 다음과 같아짐.

IF　　A는 X다
THEN　B는 Y다 $\{cf\ 0.7\}$
　　　B는 Z다 $\{cf\ 0.2\}$

❖ 나머지 10%의 의미

- 가설에 따른 규칙

$$\begin{array}{ll} \text{IF} & A\text{는 } X\text{다} \\ \text{THEN} & B\text{는 } Y\text{다 } \{cf\ 0.7\} \\ & B\text{는 } Z\text{다 } \{cf\ 0.2\} \end{array}$$

- 10%의 의미
  - A에 X 값이 주어졌다고 가정하면, B에 Y가 들어갈 확률은 시행횟수의 70%, Z가 들어갈 확률은 시행횟수의 20%라는 것을 의미
  - 나머지 10% 만큼은 어떤 값이든 들어갈 수 있음.
  - 이런 방식으로 전문가는 객체B에 알려진 두 가지 값(Y,Z) 뿐만 아니라 아직 관찰하지 못한 다른 값이 들어갈 수 있다는 가능성을 남겨둠.
  - 확신도가 전파되려면 규칙의 전건에 있는 증거가 불확실할 때 그 규칙의 후건에 대해 최종적인 확실성을 정할 수 있어야 함.
  - 전건이 하나인 규칙의 최종적인 확실성, *cf* (H,E)는 전건의 확신도 *cf*(E)에 규칙의 확신도 *cf* 를 곱하여 구할 수 있음.

$$cf\,(H,E) = cf\,(E) \times cf \qquad (3.32)$$

❖ 나머지 10%의 의미

- 다음의 예를 생각해보자.

IF 하늘이 맑다
THEN 일기예보는 화창한 날씨다 {cf 0.8}

- 하늘이 맑다의 현재 확신도가0.5라고 하면 다음과 같음.

$$cf(H,E) = 0.5 \times 0.8 = 0.4$$

- [표3-4]에 따라, 결론은 '어쩌면 화창한 날씨일 것이다'로 해석됨.

[표 3-4] 불확실함을 나타내는 용어들과 그 해석

용어	확신도
확실히 아님	−1.0
거의 확실히 아님	−0.8
아마 아닐 것임	−0.6
어쩌면 아닐 수도 있음	−0.4
알려져 있지 않음	−0.2~+0.2
어쩌면	+0.4
아마	+0.6
거의 확실히	+0.8
확실히	+1.0

❖ 전문가 시스템에서의 확신도 결정

- 여러 개의 전건을 가진 규칙의 확신도일 경우
- 다음과 같이 논리곱으로 연결된 규칙을 보자.

IF 〈증거 E_1〉
AND 〈증거 E_2〉
·
·
·
AND 〈증거 E_n〉
THEN 〈가설 H〉 {cf}

- 후건의 최종적인 확실성, 즉 가설H의 확실성은 다음과 같이 정해짐.

$$cf(H, E_1 \cap E_2 \cap \ldots \cap E_n) = min[cf(E_1), cf(E_2), \ldots, cf(E_n)] \times cf \qquad (3.33)$$

- 다음과 같은 예를 생각해보자.

IF 하늘이 맑다
AND 일기예보는 화창한 날씨다
THEN 취해야 할 행동은 '선글라스를 끼는 것' 이다 {cf 0.8}

- 하늘이 맑다의 확실성은0.9, 일기예보는 화창한 날씨다의 확실성이0.7이라면 다음과 같이 된다.

$$cf(H, E_1 \cap E_2) = min[0.9, 0.7] \times 0.8 = 0.7 \times 0.8 = 0.56$$

- [표3-4]에 따라, 결론은 '아마 오늘은 선글라스를 끼는 것이 좋을 것이다'로 해석됨.
- 다음과 같이 논리합으로 연결된 규칙을 보자.

IF ⟨증거 E_1⟩
AND ⟨증거 E_2⟩
.
.
.
OR ⟨증거 E_n⟩
THEN ⟨가설 H⟩ $\{cf\}$

- 가설 H의 확실성은 다음과 같이 결정됨.

$$cf(H, E_1 \cup E_2 \cup \ldots \cup E_n) = max[cf(E_1), cf(E_2), \ldots, cf(E_n)] \times cf \qquad (3.34)$$

$$cf(H, E_1 \cup E_2) = max[0.6, 0.8] \times 0.9 = 0.8 \times 0.9 = 0.72$$

06_확신도 이론과 증거 추론

- 다음 예를 보면

IF 하늘이 흐리다
OR 일기예보는 비오는 날씨다
THEN 취해야 할 행동은 '우산을 가져가는 것' 이다 $\{cf\ 0.9\}$

- '하늘이 흐리다'의 확신도는 0.6, 일기예보는 '비오는 날씨다'의 확신도가 0.8이라고 하면 다음과 같이 됨.

$$cf\,(H, E_1 \cup E_2) = max\,[0.6, 0.8] \times 0.9 = 0.8 \times 0.9 = 0.72$$

- 결론은 '오늘은 거의 확실히 우산을 가져가야 한다'로 해석됨.

06_확신도 이론과 증거 추론

가끔 두 개 또는 그 이상의 규칙이 같은 가설에 영향을 줄 수 있다. 전문가 시스템은 이 상황을 어떻게 처리하는가

- 두 개 또는 그보다 많은 규칙을 수행한 결과로 같은 결론을 얻으면, 가설 하나의 확신도를 결합하여 제공하도록 이들 규칙에 대한 각각의 확신도를 합쳐야 함.
- 기반 지식은 다음의 규칙으로 이루어져 있다고 가정하자

규칙 1

IF A는 X다

THEN C는 Z다 $\{cf\ 0.8\}$

규칙 2

IF B는 Y다

THEN C는 Z다 $\{cf\ 0.6\}$

- 규칙 1과 규칙 2가 점화되었을 경우, Z 값을 갖는 객체C에 어떤 확신도 값이 들어가야 할까?
- 상식적으로는 서로 다른 출처(규칙 1과 규칙 2)에서 같은 가설(C는Z다)을 지지하는 두 개의 증거(A는 X다와 B는 Y다)가 있다면, 이 가설에 대한 신뢰도는 증가하고, 이는 하나의 증거만 관찰했을 때보다 더 강해짐.

- 결합된 확신도를 계산하려면 다음 식을 사용한다.

$$cf(cf_1, cf_2) = \begin{cases} cf_1 + cf_2 \times (1 - cf_1) & \text{if } cf_1 > 0 \text{ and } cf_2 > 0 \\ \dfrac{cf_1 + cf_2}{1 - min[|cf_1|, |cf_2|]} & \text{if } cf_1 < 0 \text{ or } cf_2 < 0 \\ cf_1 + cf_2 \times (1 + cf_1) & \text{if } cf_1 < 0 \text{ and } cf_2 < 0 \end{cases} \quad (3.35)$$

  - 여기서 쓰인 기호는 다음과 같다.
    - cf_1: 규칙 1에 의해 정해진 가설 H에 대한 확신
    - cf_2: 규칙 2에 의해 정해진 가설 H에 대한 확신
    - $|cf_1|$과 $|cf_2|$: cf_1과 cf_2의 절대값

- 다음과 같이 가정해보자.

$$cf(E_1) = cf(E_2) = 1.0$$

  - (3.32)에서 다음을 얻는다.

$$cf_1(H, E_1) = cf(E_1) \times cf_1 = 1.0 \times 0.8 = 0.8$$
$$cf_2(H, E_2) = cf(E_2) \times cf_2 = 1.0 \times 0.6 = 0.6$$

  - 또한 (3.35)에서 다음을 얻는다.

$$cf(cf_1, cf_2) = cf_1(H, E_1) + cf_2(H, E_2) \times [1 - cf_1(H, E_1)]$$
$$= 0.8 + 0.6 \times (1 - 0.8) = 0.92$$

- 이 예는 가설에 대해 신뢰가 점점 증가하는 것을 보여 준다.

- 이번에는 규칙의 확신도가 반대 부호인 경우를 생각해보자.
- 다음과 같이 가정하면,

$$cf(E_1) = cf(E_2) = -1.0$$

  - 다음과 같이 된다.

$$cf_1(H, E_1) = 1.0 \times 0.8 = 0.8$$
$$cf_2(H, E_2) = -1.0 \times 0.6 = -0.6$$

  - (3.35)에서 다음을 얻는다.

$$cf(cf_1, cf_2) = \frac{cf_1(H, E_1) + cf_2(H, E_2)}{1 - min\,[|cf_1(H, E_1)|, |cf_2(H, E_2)|]} = \frac{0.8 - 0.6}{1 - min\,[0.8, 0.6]} = 0.5$$

- 이 예는 하나의 규칙(규칙 1)은 가설을 지지하지만, 다른 하나의 규칙(규칙 2)은 그렇지 않을 때 결합된 확신도, 즉 최종적인 신뢰를 얻는 방법을 보여줌.

- 규칙의 확신도가 음의 부호를 갖는 경우를 생각해보자.
- 다음과 같이 가정하면,

$$cf(E_1) = cf(E_2) = -1.0$$

- 그러면 다음과 같다.

$$cf_1(H, E_1) = -1.0 \times 0.8 = -0.8$$
$$cf_2(H, E_2) = -1.0 \times 0.6 = -0.6$$

- (3.35)에서 다음을 얻는다.

$$\begin{aligned} cf(cf_1, cf_2) &= cf_1(H, E_1) + cf_2(H, E_2) \times [1 + cf_1(H, E_1)] \\ &= -0.8 - 0.6 \times (1 - 0.8) = -0.92 \end{aligned}$$

- 이 예는 어떤 가설에 대해 불신이 점점 증가하는 것을 나타냄.
- 확신도 이론은 베이즈 추론보다 훨씬 실용적인 대안
- 확신도를 결합하는 휴리스틱 방식은 확률 값을 결합하는 방식과는 다름.
- 확신도는 '수학적으로 완전' 하지는 않지만, 인간 전문가의 사고 방식을 모방한 것임.

07_FORECAST: 확신도의 응용 사례

- 전문가 시스템은 내일 비가 올지 안 올지를 예측해야 한다. 즉, 값이 여러 개가 될 수 있는 객체 내일에 대한 확신도를 설정해야 한다.
- 작업을 간단하게 하기 위해 앞에서 사용한 규칙 집합을 사용하자.

- 기반 지식

/* FORECAST: 확신도의 응용

control *cf*

control '임계값 0.01'

규칙 1

IF 오늘은 비가 온다

THEN 내일은 비가 온다 {cf 0.5}

규칙 2

IF 오늘은 맑다

THEN 내일은 맑다 {cf 0.5}

규칙 3

IF 오늘은 비가 온다

AND 강우량이 적다

THEN 내일은 맑다 {cf 0.6}

규칙 4

IF 오늘은 비가 온다

AND 강우량이 적다

AND 기온이 낮다

THEN 내일은 맑다 {cf 0.7}

규칙 5

IF 오늘은 맑다

AND 기온이 높다

THEN 내일은 비가 온다 {cf 0.65}

규칙 6

IF 오늘은 맑다

AND 기온이 높다

AND 하늘이 흐리다

THEN 내일은 비가 온다 {cf 0.55}

seek 내일

❖ 확신도 기반 전문가 시스템: 대화

- 확신도에 기반을 둔 부정확한 추론 기법을 적용하려면 전문가 시스템은 사용자에게 객체 값뿐만 아니라 그 값과 관련된 확신도를 함께 입력 받아야 한다.
- 예를 들어, Leonardo 척도를 0~1.0 사이의 값을 사용하면 다음과 같은 대화가 가능하다.

- 오늘의 날씨는 어떻습니까? ⇒ 비가 온다

규칙 1

IF 오늘은 비가 온다

THEN 내일은 비가 온다 {cf 0.5}

- cf(내일은 비가 온다, 오늘은 비가 온다)=cf(오늘은 비가 온다)×cf=1.0×0.5= 0.5
- 내일은 비가 온다 {0.50}

- 오늘의 강우량은 어떻습니까?⇒ 적다

- 강우량이 적다고 어느 정도로 믿고 있습니까? 0~1.0 사이의 숫자로 확실성을 입력해주십시오. ⇒ 0.8

07_FORECAST: 확신도의 응용 사례

규칙 3

IF	오늘은 비가 온다
AND	강우량이 적다
THEN	내일은 맑다 {cf 0.6}

- cf(내일은 맑다, 오늘은 비가 온다 ∩ 강우량이 적다)
  = min[(오늘은 비가 온다), cf(강우량이 적다)] × cf = min[1, 0.8] × 0.6 = 0.48

- 내일은 비가 온다 {0.50}
- 내일은 맑다 {0.48}

- 오늘 기온은 어떻습니까? ⇒ 낮다
- 기온이 낮다고 어느 정도로 믿고 있습니까? 0~1.0 사이의 숫자로 확실성을 입력해주십시오. ⇒ 0.9

규칙 4

IF	오늘은 비가 온다
AND	강우량이 적다
AND	기온이 낮다
THEN	내일은 맑다 {cf 0.7}

07_FORECAST: 확신도의 응용 사례

- cf(내일은 맑다, 오늘은 비가 온다 ∩ 강우량이 적다 ∩ 기온이 낮다)
  = min[cf(오늘은 비가 온다), cf(강우량이 적다), cf(기온이 낮다)] × cf
  = min[1, 0.8, 0.9] × 0.7 = 0.56

- 내일은 맑다 {0.56}
- 내일은 비가 온다 {0.50}

- cf(cf(Rule:3), cf(Rule:4)) = cf(Rule:3) + cf(Rule:4) × (1 – (cfRule:3))
  = 0.48 + 0.56 × (1 – 0.48) = 0.77

- 내일은 맑다 {0.77}
- 내일은 비가 온다 {0.50}

- 이제 내일의 날씨가 맑을 가능성은 거의 확실하다고 결론지을 수 있다. 또한 비가 약간은 내릴 수도 있다.

❖ 베이즈 추론

- 베이즈 기법을 적용한 전문가 시스템인 PROSPECTOR는 광석 매장소를 찾는 탐사 지리학자를 돕기 위해 개발됨.
  - 이 시스템은 매우 성공적이었는데, 일례로 PROSPECTOR는 지리학적, 지구 물리학적, 지구 화학적 데이터를 사용하여 워싱턴 주에 있는 톨만산 근처에 몰리브덴이 있다는 사실을 예측
  - PROSPECTOR 팀은 잘 알려진 광물 매장소에 대한 데이터와 믿을만한 통계 정보를 사용할 수 있었고, 각 사건의 확률 역시 잘 정의되어 있었음.
  - 또한 베이즈 접근법을 적용하기 위한 제약 조건인 증거의 조건부 독립성도 가정할 수 있었음.

- 전문가 시스템을 적용할 수 있는 여러 분야에는 믿을만한 통계 정보가 없거나 증거의 조건부 독립을 가정할 수 없었음.
  - 그 결과, 많은 연구자는 베이즈 방법이 자신의 작업에 적합하지 않다는 사실을 알게 됨.
  - 쇼틀리프와 뷰캐넌은 의학 분야에서 필요한 데이터를 얻을 수 없었기 때문에 MYCIN에 고전적인 확률 기법을 사용할 수 없었음.

08_베이즈 추론과 확신도의 비교

❖ 확신도 접근법

- 확신도 접근 방식은 확률 이론에서 볼 수 있는 수학적 정확성이 부족했지만, 진단, 특히 의학 진단분야에서는 베이즈 추론을 능가하는 것으로 드러남.
- MYCIN과 같은 진단 전문가 시스템에서규칙과 확신도는 전문가의 지식과 직관적 판단에서 비롯한다. 확신도는 확률이 알려지지 않거나 확률 값을 얻는 데 드는 비용이 매우 클 때 사용.
- 증거 추론 메커니즘은 신뢰의 정도가 서로다른 증거를 다룰 수 있을 뿐만 아니라 점진적으로 습득하는 증거, 가설의 논리적 결합도 다룰 수 있음.
- 게다가 확신도 접근은 규칙기반 전문가 시스템에서 제어의 흐름을 훨씬 잘 설명할 수 있음.

08_베이즈 추론과 확신도의 비교

❖ 베이즈 추론 VS 확신도

- 베이즈 접근과 확신도는 서로 다르지만, 개인적이고 주관적이며 질이 좋은 정보의 양을 잴 수 있는 전문가를 찾아야 한다는 공통된 문제점이 있음.
- 인간은 편협해지기 쉽다. 그러므로 불확실성을 다루는 기법을 선택할 때는 주제 전문가에게 전적으로 맡김.
- 베이즈 방법은 신뢰성 있는 통계 데이터와 진행을 이끌어갈 지식 공학자가 있고, 결정에 대해 전문가와 분석적이고 진지한 대화를 할 수 있는 경우에 가장 적합함.
- 위 조건 중 하나라도 부족하다면 베이즈 접근법은 너무 멋대로가 되고, 심지어 편협해져서 의미있는 결과를 얻을 수 없음.
- 또한 베이즈 신뢰 전파는 지수적으로 복잡하여 큰 기반 지식에는 실용적이지 않음.
- 확신도 기법은 비록 이론적인 기반이 부족하지만, 전문가 시스템에서 불확실성을 다룰 수 있는 단순한 접근 방식을 제공하고, 여러 응용 분야에서 인정할 만한 결과를 제시

❖ 규칙기반 전문가 시스템에서 불확실성 다루기 요약

- 불확실성은 믿을만한 결론을 위한 정확한 정보가 부족하다는 것을 뜻함. 전문가 시스템에서 지식이 불확실해지는 이유는 상관관계가 취약한 함축, 부정확한 언어, 데이터의 분실, 다른 전문가의 관점을 통합하면서 겪는 어려움 때문이다.

- 베이즈 규칙
  - 확률 이론은 전문가 시스템에서 불확실성을 다룰 수 있도록 정확하고 수학적으로 올바른 접근법을 제시한다. (베이즈 규칙)
  - 광물 탐사용 전문가 시스템인 PROSPECTOR는 시스템 전체에 걸쳐 불확실성을 전파하는 증거의 베이즈 규칙을 적용하여 성공한 첫 시스템이다.
  - 베이즈 접근법에서 가설 H의 사전 확률과 증거 E가 있는 경우, 신뢰도를 측정 시 충분 가능성 LS 값을, 증거가 없는 경우 가설에 대한 불신도를 측정 시 필요 가능성 LN 값을 제공해야 한다.

$$\begin{array}{ll} \text{IF} & E\text{가 참 } \{LS, LN\} \\ \text{THEN} & H\text{가 참 } \{\text{사전 확률}\} \end{array}$$

  - 베이즈 접근법을 적용 시 **증거의 조건부 독립성**, **믿을 만한 통계 데이터**, 각 가설에 대해 **사전 확률**이 필요하다.

09_요약

- 확신도 이론
  - 확신도 이론은 베이즈 추론의 대안으로 많이 쓰인다. 이 이론의 기본 원칙은 진단 의학 전문가 시스템인 MYCIN에 도입되었다.
  - 확신도 이론은 전문가 시스템에서 불확실성을 다루는 판단적 접근 방식을 제공한다. 전문가는 증거 E를 관찰했을 때 가설H에 대한 신뢰 수준을 표현하기 위해 확신도 cf를 제공해야 한다.

$$\begin{array}{ll} \text{IF} & E\text{가 참} \\ \text{THEN} & H\text{가 참 } \{cf\} \end{array}$$

  - 확신도는 확률이 알려지지 않았거나 확률 값을 쉽게 얻을 수 없는 경우에 사용한다. 확신도 이론은 신뢰 정도가 서로 다른 증거뿐만 아니라 점진적으로 습득하는 증거, 가설의 논리적 결합을 다룰 수 있다.

- 베이즈 접근과 확신도는 서로 다르지만, 개인적이고 주관적이며 질이 좋은 정보의 크기를 잴 수있는 전문가를 찾아야 한다는 공통된 문제점이 있다.

Thank You !