Historia Mathematatica Vol. 14, No. 2, 2001, 93-100

수학적 대상으로서 '애매모호'에 대한 고찰

서경대학교 **박창균**

Abstract

The problem of vagueness has been investigated for a long time by philosophers and mathematicians. There are three approaches in mathematics to the problem, which are probability theory, fuzzy logic, and rough set theory. In this paper I introduce these theories and their meanings.

0. 서론

일반적으로 수학을 포함한 모든 학문에서 불확실하거나 애매모호한 대상은 환영을 받는 위치에 있지 않다. 특히 엄밀한 학문의 전형으로 인식되어 온 수학에서는 더욱 더 그러하다. 수학에서의 이러한 인식에는 그리스 기하학이 주었던 깊은 인상과 현대 수학의 공리적이고 연역적 체계에서 비롯된 느낌이 크게 작용한 듯 하다. 그래서 많은 사람들은 수학과 불확실한 것 또는 수학과 애매모호한 것은 별로 관계가 없다고 생각한다. 그런데 사실 우리 주변에는 확실하고 명확한 것보다는 불확실하고 애매모호한 것들이 더 많다. 이러한 분야를 학문적 대상에서 제외하는 것은 학문의 목적을 다시 묻게 하고, 자칫 학문을 관념의 유희 내지 하나의 '게임'으로 전락시킬 위험이 있다. 그렇다고 학문이 현실과 완전 밀착해야 한다고 주장하거나 추상적인 면을 학문에서 완전히 배제해야 한다고 주장하는 것은 아니다. 그리고 학문에 어느 정도 게임의 성격이 있는 것도 사실이지만 애매모호한 것들을 탐구 대상에서 제외하는 것은 학문이 삶과 완전히 유리되어 있지 않는 한 학문 그 자체로 손실이 크다. 그러나 수학을 포함한 제 학문들은 표면적으로는 애매모호한 것들을 거부하고 배격했지만, 실상은 불확실성과 애매모호성을 상당히 의식하고 그것들을 해소하려는 노력을 경주했고 또 그것이 학문이 걸어온 발자취였다고 해도 과언이 아니다. 여기서 구분해야 할 것은 불확실하고 애매모호한 것을 대상으로 다루는 것과 학문 자체가 가지는 성격이다. 즉 수학을 포함한 제 학문은 그 자체는 애매모호하지 않아야 하지만 불확실한 것들과 애매모호한 것들은 얼마든지 학문적 대상으로 할 수 있고 또 그렇게 해야만 하는 당위성도 있다.

수학은 모든 학문의 언어로서 자연현상이나 사회현상을 기술하는 데 유용한 도구로서 역할을 담당해왔다. 그래서 혹자는 수학을 학문의 시녀로 혹자는 학문의 여왕으로 묘사했다. 수학이 모든 학문의 시녀이든 여왕이든 분명한 것은 수학이 담당하는 언어적 기능이다. 언어적 기능은 다른 학문들이 묘사하는 것을 돕는 시녀가 되기도 하고, 수학 없이 그 대상들을 효과적으로 다룰 수 없기에 여왕의 역할을 동시에 가능하게 한다. 따라서 새로운 주제들을 접근하는데 있어서 가장 필요한 것은 그것들을 수학적으로 정형화하도록 새로운 수학적 언어(개념)를 개발하는 일이다. 수학적 언어가 개발되어야 비로소 그 주제들을 수학적으로 취급할 수 있게 된다. 집합론은 그것이 가진 포괄성과 일반성 때문에 수학적 언어의 기초가 된다. 좀 강하게 표현하여 모든 과학이 수학적 언어로 표현되어야 하고 집합론이 수학의 기본 어휘라면, 과학이나 수학에서 기존 체계를 뛰어 넘는 새로운 개념을 도입하려고 할 때 새로운 집합론이 제시 되어야함은 당연한 일이다. 이 글에서 앞으로 불확실하고 애매모호한 것들-그것이 언어로 표현된 실체이든 속성이든 간에-을 '애매모호'로 나타내고, 특별한 언급이 없다면 '애매모호'와 '애매'를 구별없이 혼용하기로 한다.

'애매모호'에 대한 수학적 접근은 비교적 짧은 역사를 가지고 있다. 불확실성을 다루는 확률론은 그 중 가장 일찍 시작되었지만 다른 수학의 역사에 비하면 결코 길다고 할 수 없다. 물론 사람들이 확률에 대하여 생각한 것은 꽤 거슬러 올라가겠지만 체계적이라고 보여지는 접근이 이루어 진 것은 근대 이후이다. 언어나 정보에 대한 '애매모호'를 다루게 된 것은 아주 최근의 일이다. 퍼지이론과 러프집합론(rough set theory)은 이러한 '애매모호'를 다루는 대표적인 것들인데, 이것들은 특히 컴퓨터의 발전과 관계가 있으며 그동안 학문이 가졌던 외연을 확장 할 수밖에 없는 사회적이고 문화적 요인들이 작용했다고 보인다.

본 논문은 '애매모호'에 대한 수학적 접근들을 소개하고 이에 대해 철학적으로 음미하려는 데에 그 목적이 있다. 이를 위해 먼저 '애매모호'를 다루는 이론- 확률론, 퍼지논리, 러프집합론 등-을 역사적 배경과 함께 소개하고, 이중에 특히 러프집합론을 중심으로 논의를 전개한다.

1. '애매모호'의 양상과 확률론

우리가 '애매모호'하다고 할 때 그것은 단순히 어느 하나의 현상에 고착되어 있는 것은 아니다. '애매모호'의 양상들은 매우 다양하며 다음과 같이 정리될 수 있다[2, 36].

(1) 지식이 부족한 경우(incomplete)

(2) 해석이 여러 가지인 경우(ambiguity)

(3) 미래의 일인 경우(randomness)

(4) 부정확한 경우(imprecision)

(5) 정의할 수 없거나 정의해도 무의미한 경우(fuzziness)

(1)의 경우는 지식이 충족되면 해결되는 애매성이다. 즉 아랍어를 전혀 모르는 사람이 아랍어를 들을 때 비록 의미가 있는 내용일 지라도 지식부족으로 인하여 발생하는 애매성이다. (2)는 '다리'라 할 때 신체의 일부로서의 다리인지 교량을 의미하는지 불분명한 경우를 말한다. 즉 어느 하나가 여러 가지 중에서 어디에 속하는지가 불분명한 경우이다. (3)은 내일 비가 올 것인지, 주사위를 지금 던지면 3의 눈이 나올 것인지 등 일어날 일에 대한 불확실성이 주는 애매성이다. (4)는 오류나 잡음이 개재되어 생기는 부정확에서 오는 애매성이다. (5)는 언어가 가지는 애매성이다. 즉 아름답다거나 늙었다고 할 때 객관적으로 정의하기가 쉽지 않다. 이러한 언어적 표현에는 누구나 동의하는 경우보다는 애매한 경우가 많을 것이다. 혹자는 애매와 모호를 구분하여 (2)를 애매하다고 하고, (5)를 모호하다고 하기도 한다.

그런데 이러한 애매의 양상 중에서 먼저 수학적 관심의 대상이 된 것은 (3)의 경우이다. 확률론이 다루는 대상이 바로 이것이라 할 수 있다. 예로부터 여러 가지 게임과 도박이 있었지만 15세기 말까지 확률에 대한 수학적인 접근이 이루어졌다고 볼 수는 없다. 확률론의 기원이 되었다고 여겨지는 문제는 소위 '득점의 문제' 인데, 이 문제는 1494년 파촐리(Pacioli)에 의해 처음으로 수학적 연구 대상으로 그의 저서 *Summa*에 제시되었다. 그리고 1654년은 확률론에 있어서 아주 뜻깊은 해였다. 이 해에 파스칼과 페르마는 서신교환을 통하여 확률론의 기본원리들을 형성했다. 1657년 호이겐스는 파스칼과 페르마의 작업에 기초해서 확률론의 첫 논문이라고 할 수 있는 "우연게임에서의 추론에 관하여"를 발표하였다. 그 후 그론트(Graunt)는 보험을 위한 확률에 대한 작업을 했고 이것은 수리통계의 시작을 알리는 것이었다. 지금은 표준으로 되어있는 순열과 조합의 결과를 포함하는 확률론의 첫 주요 업적과 자연수를 지수로 가지는 이항정리에 대한 첫 번째 증명 그리고 '대수의 법칙' 등은 1713년 야콥 베르누이에 의해 성취되었다. 그 후 1770년대에 라플라스가 **확률의 해석적 이론**에서 확률에 대한 수학적 이론을 우연의 게임을 넘어서는 영역으로 확장시켰는데, 확률론이 공리적으로 완성된 것은 20세기 초 콜모고로프에 의해서였다.

확률의 개념에 대해서는 여러 가지 입장이 존재한다. 위에서 언급한 라플라스는 나타날 수 있는 모든 경우의 수에 대하여 확률을 구하고 있는 사건이 나타나는 경우의 수의 비율이 확률이라고 했다[5, 6-7]. 이러한 확률 개념의 구체적인 예는 주사위를 던졌을 때 3의 눈이 나올 확률과 같은 것이다. 이런 의미의 확률은 경험적 관찰에 근거했다기보다는 선험적인 지식에 토대를 두고 있고, 사물이 가지는 객관적 성질을 반영하는 것이 아닌 베르누이의 표현대로 확실성의 도수[3, 48] 내지 합리적 기대치이다. 그러나 이런 입장에서 어떤 사건의 확률은 그 사건이 실제로 일어나는 빈도를 나타낸다고 볼 수는 없고, 동등한 실현성을 알지 못하는 경우에 대해서 동등한 확률을 부여하는 소위 '무차별의 원리'에 호소했다는 비판이

있다. 또 하나의 확률에 대한 입장은 '빈도설'이다. 빈도설에서는 확률을 상관빈도의 극한값이라고 규정한다. 이 입장은 라이헨바흐의 다음과 같은 지적에 잘 나타나 있다.

> "확률진술은 반복되는 사건의 상관빈도 즉 총합의 백분율로 계산되는 빈도를 나타낸다. 그런데 이것은 과거에 관찰된 빈도에서 나오는 것이고 미래에도 이와 비슷한 빈도로 일어난다는 가정을 포함한다[9, 236]."

이 입장은 라플라스가 생각했던 선험적인 확률과 달리 경험적인 정보를 전제하고, 실재적인 성격을 갖는다. 그러나 이 입장은 이미 알고 있는 것을 확대하는데 따른 부담을 안게 된다. 그러한 확대는 어떻게 정당화될 수 있는가? 이 문제는 과학의 방법론으로서 귀납법이 어떻게 정당화되는가? 하는 과학철학적 논의를 요구한다. 귀납법을 과학적 방법론으로 받아들이기를 거부하는 과학철학자도 있으나 라이헨바흐는 귀납법의 '실용적' 정당화를 주장한다. 확률에 대한 마지막 입장을 소개한다면, 확률이 어떤 가설을 확증하는 정도를 나타낸다고 하는 것이다. 이런 입장을 '귀납설'이라고 하는데 이를 대표하는 카르납은 확률로써 증거에 의거한 가설에 주어진 강도, 즉 확증정도를 결정한다고 주장한다[4, 270]. 이 입장은 빈도설처럼 경험적이고 실제적이 아니라 논리적이고 증거 의존적이다. 따라서 이러한 견해는 실세로 일어나는 어떤 사건을 예측하는 도수로서의 확률이 아닌 과학적 지식을 판단하고 확립하는데 요구되는 확률을 의미한다.

2. 퍼지논리와 러프집합론

확률론이 상대적으로 오랜 수학적 전통 속에서 이루진 애매에 대한 접근 방법인데 비해 퍼지논리와 러프집합론은 20세기 후반에야 대두되었고 수학 내적인 전통에서라기보다 수학 외적인 요구에 의해 전문 수학자가 아닌 사람들에 의해 제기되었다. 이미 언급한 것과 같이 수학에서는 확률론에서 취급하는 애매성(randomness)을 제외한 어떠한 애매성도 수학의 대상으로 본격적으로 부상할 수 없었다. 이러한 분위기는 순수수학을 하는 사람들에게는 여전히 유지되고 있다고 보여진다. 그런데 컴퓨터가 발달하고 이에 따라 불확실하고 애매한 정보를 처리해야하는 압력이 커지면서 애매성에 대한 관심이 높아졌다. 전술한 바와 같이 새로운 주제들을 취급하기 위해서는 그것을 형식화하는 어휘가 존재해야하고 가장 기본적인 어휘를 정의하는 새로운 집합론이 요구되었다. 이렇게 하여 등장한 것이 퍼지집합론과 러프집합론이다. 따라서 이 집합론들을 '정보화 시대의 집합론'이라고 부를 수 있다.

퍼지집합의 개념은 1965년 자데(Zadeh)에 의해 제안되었다. 그는 엄격한 제어 공학자였지만 이치논리의 한계를 의식했던 것 같다. 집합론은 칸토어에 의해 19세기 말에 제안되었고 집합의 개념은 명확히 구분 가능한 것이었다. '젊은 사람의 모임'이라든가 '키 큰 사람의 모임'과 같이 소속을 명확하게 규정할 수 없는 것들은 집합이 될 수 없었다. 칸토어의 개념에

따르면, 만약 어떤 원소가 한 집합에 소속되면 1의 값을 주고 소속되지 않으면 0을 값을 주어 특성함수는 단지 두 가지 값만 가질 수 있다. 그런데 자데는 한 원소가 어떤 집합에 소속되는 정도를 0과 1을 포함한 0과 1 사이의 값을 갖도록 확장한 것이다. 따라서 상기한 '젊은 사람의 모임'이나 '키 큰 사람의 모임'도 집합(퍼지)이 되고 새로운 결과를 낳게 된다. 즉 30살인 사람은 '젊은 사람의 모임'에는 0.8정도 들어간다면, '젊지 않은 사람의 모임'에 0.2정도 들어간다고 할 수 있어서 논리학의 가장 기본 법칙인 모순율과 배중률이 성립되지 않게 된다. 퍼지논리는 자연언어에 대하여 수학적 접근을 가능하게 하는 인공지능의 가장 중요한 도구의 하나로 등장하게 되었다. 문제는 소속함수의 값을 어떻게 결정할 수 있느냐는 것이다. 언어가 가지는 애매성 때문에 사람에 따라 다를 소속정도는 객관성을 갖기 어려울 것이라는 반론이다. 그러나 일단 소속정도가 정의되고 규칙이 결정되면 퍼지논리는 효율성과 실용성에서 상당한 위력을 발휘하게 된다. 그 결과 오늘날 가전제품에 퍼지논리를 구현한 것들을 많이 찾아 볼 수 있다.

러프집합은 1982년 파블락(Pawlak)이 제안한 집합론으로서 애매에 대한 새로운 수학적 접근 방식이다. 러프집합론이 가지고 있는 기본전제는 언어로 말하여 지거나 표현된 모든 대상들은 어떤 정보-자료, 지식 등-와 연관되어져 있다는 것이다. 예를 들어 대상들이 어떤 질병을 앓고 있는 환자들이라면, 그 질병에 대한 증상들이 그 환자에 대한 정보를 형성한다. 같은 정보에 의해 특성화된 대상들은 그들에게 허용된 정보의 관점에서 '식별 불가능하다'고 하고, 이 '식별 불가능한 관계'는 러프집합론의 수학적 기초를 이룬다. 모든 식별 불가능한 대상들의 집합들을 '기초집합'이라 하고 그 세계에 관한 지식의 기본적 알갱이를 형성한다. 기초집합의 합집합은 보통집합(명확한 집합)이고, 기초집합의 합집합이 아니라면 러프집합이 된다. 즉 러프집합은 어떤 기초집합의 일부만을 포함하는 경계에 걸쳐있는 집합이며, 어떤 집합에 확실하게 들어가지도 않고 그 집합의 여집합에도 확실하게 들어가지 않는 원소를 가지고 있다. 이런 경계에 걸쳐 있는 원소를 '경계원소'라 한다면, 러프집합은 경계원소를 반드시 포함하는 집합이라고 할 수 있다.

러프집합에서는 지식을 '알갱이 구조'로 파악한다. 따라서 어떤 대상들은 구별되지 아니하고 같은 것으로, 즉 같은 알갱이로 간주한다. 애매한 개념은 명석 판명한 개념과는 달리 그들이 포함하고 있는 원소들에 의해 특성화 될 수 없다. 그러면 이러한 애매한 개념들은 어떻게 다룰 것인가? 러프집합에서는 애매한 개념을 그 개념의 '상한근사(upper approximation)'와 '하한근사(lower approximation)'를 통하여 다룬다. 상한근사와 하한근사는 애매한 개념이 아니라 정확한 개념이다. 상한근사는 그 개념에 '확실하게' 속하는 모든 대상들의 집합(기초집합)이고, 하한근사는 그 개념에 속하는 것이 '가능한' 모든 기초집합들을 의미한다. 따라서 상한근사와 하한근사의 차이에 해당하는 것이 '경계영역'이 된다. 근사는 러프집합론의 기본적인 두 연산이다. 러프집합론은 다양한 분야에서 응용되어 왔다. 인공지능과 인지과학, 특별히 기계학습 분야에서 근본적인 방법론이다. 그 외에 데이터해석, 근사분류, 화상처리 등에서 그 유용성이 입증되고 있으며 결정지원시스템(decision support system)에서 특

별한 중요성을 가지는 듯하다.

3. 맺는 말

인류의 역사가 애매성을 해소하려는 노력으로 점철되어왔음은 주지하는 바이다. 소위 과학기술의 시대라는 요즈음도 인간의 운명이나 미래를 예측하기 위하여 학문적으로 정당화될 수 없는 방법들이 우리 주위에서 쉽게 목격된다.

수학에서 애매성을 취급하게 된 것은 수학이라는 학문의 특성상 다른 수학 분야보다는 일천하다. 확률론은 그나마 오래된 수학적 접근이지만 언어의 애매성과 관련된 퍼지논리나 러프집합론은 최근에 대두되었다. 이들 집합론이 등장하게 된 데에는 컴퓨터와 연관이 깊다. 컴퓨터는 지식을 저장하고 활용하는 수단이다. 그러나 문제는 지식이 무엇인지 이해하고, 지식을 표현하는 방법을 찾고 저장된 지식으로부터 유용한 정보를 추출해 내기 위하여 기계가 단순히 '기계'로서 존재해서는 안 된다는 것이다. 인간의 지식은 자연언어로 표현된다. 이 지식을 컴퓨터가 자동적으로 처리하기 위해서는 자연언어가 가지는 애매성이 문제가 된다. 이런 이유로 퍼지논리와 러프집합론이 등장했다고 볼 수 있다. 언어의 애매성을 다루는데 있어서 퍼지논리는 소속정도를 정의함으로써, 그리고 러프이론은 상한근사와 하한근사라는 두 명확한 집합으로써 접근한다. 러프이론의 주된 이점은 데이터에 대한 사전 또는 사후의 추가적 정보가 필요 없다는 것이다.

언어가 가지는 애매성을 해소하려고 하는 퍼지논리와 러프집합론의 철학적 함축은 무엇일까? 현대 철학에서 언어는 중심적인 주제이다. 분석철학에서도 그러하고 해석학의 원천도 언어이다. 이때의 언어는 인공언어가 아닌 일상적인 자연언어이다. 자연언어 속에 인간의 삶과 문화가 용해되어 있고 또 그 언어가 가지는 삶의 형식에 근거해서 인식하고 언어와 대상을 연결짓는다. 애매한 것들에 대한 접근은 결국 인식과 이해의 문제로 귀착된다. 언어행위는 삶을 규정하는 형식이고 모든 이해의 전제조건이면서 동시에 제약조건이다[1, 239]. 퍼지집합의 소속함수의 값을 결정하는 것과 식별 불가능한 관계에 의한 분류가 퍼지논리와 러프집합론의 가장 중요한 일이 라면, 이것은 과연 어떻게 가능한 것인가? 어느 정도 상대적일 수밖에 없다는 사실을 인정한다고 해도 그래도 어떤 범위에서 오차 내에 있다는 것은 언어의 역사성과 사회성에 때문이라고 생각할 수 있다. 후기 비트겐슈타인의 '삶의 형식'이나 후기 후설의 '생활세계'라는 개념은 소속함수의 값을 결정하고 식별 불가능한 관계에 의한 분류를 지지해주는 기반이 된다고 볼 수 있다.

19세기 말과 20세기 초에 걸쳐 이루어진 위대한 과학적 성과들의 일부를 부정적인 진술을 통해 나타낸다면 다음과 같다.

- 열역학 제2 법칙: 엔트로피는 감소할 수 없다.

- 상대성 이론: 빛의 속도보다 빠를 수 없다.
- 하이젠베르그의 불확정성 원리: 물체의 위치와 운동량은 동시에 측정할 수 없다.
- 괴델의 불완전성 정리: 자연수 체계를 포함하는 형식체계가 무모순이면 완전하지 않다.
- 카오스 이론: 장기예측은 불가능하다.

그러면 '애매모호의 원리'와 같은 것이 있다면 그것이 함축하는 부정적 진술은 무엇일까? 만약 그 애매함이 시간이 지나도 해소되지 않는 그러한 언어의 애매라면, 그것에 접근하는 노력에도 불구하고 "애매는 취급되는 한 이미 애매하지 않다."는 것일 것이다. 애매가 대상화되어 취급되고 가공될 때 그것은 이미 애매함을 탈피하여 애매에 근사한 다른 애매한 것이 된다. 그런 의미에서 애매에 대한 수학적 접근들은 그것이 가지는 효율성과 실용성은 인정되지만, 애매 그 자체가 본원적으로 해결되었다고 할 수는 없는 것이다. 그러나 해결될 수 없는 애매함이 남아 있다는 것은 인간에게는 불행이기보다는 행복일지도 모른다.

참고 문헌

1. 박순영, "해석학적 방법론과 사회역사 연구," **해석학은 무엇인가**(한국해석학회 엮음), 지평문화사, 1995.
2. 向殿政男/전자신문 출판국 옮김, **알기쉬운 퍼지이론**, 전자신문사, 1991.
3. Carnap, R., *Logical Foundations of Probability*, University of Chicago Press, Chicago, 1951.
4. Carnap, R., "Statistical and Inductive Probability" in *The Structure of Scientific Thought* ed. by E. Madden, Cambridge, Mass., 1960.
5. Lapace, P., *A Philosophical Essay on Probabilities*, English tras. by Truscott and Emory, Dover Publications, Inc., New York, 1951.
6. Mises, R., *Probability Statistics and Truth*, George Allen and Unwin LTD, London, 1957.
7. Pawlak, Z., "Rough Set," *International Journal of Computer and Information Sciences* 11(1982), 341-356.
8. Pawlak, Z., *Rough Sets - Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1991.
9. Reihenbach, H., *The Rise of Scientific Philosophy*, University of California Press, Berkeley and Los Angeles. 1958.

10. Zadeh, L., "Fuzzy Logic and Approximated Reasoning," *Synthese* 30, 1975.
11. Zimmermann, H., *Fuzzy Set Theory and Its Applications*, second edition, Kluwer Academic Publishers, 1990.