

# 과학 응용 데이터베이스 관리를 위한 시간지원 데이터 모델<sup>+</sup>

김진호\* · 옥수호\*\*

## A Temporal Data Model for Managing Scientific Databases

J in-Ho K im · Soo-Ho Ok

### 〈요 약〉

최근 컴퓨터 응용이 우주 항공, 천체 기상, 환경 관리, 공장 자동화(FA) 등의 분야로 확대되면서 물리, 화학, 생물, 기계 등의 과학 응용에서 생성되는 자료를 처리하는 기법에 대한 요구가 늘어나고 있다. 이들 과학 응용에서는 물리적 실험 장치나 측정 기계, 또는 시뮬레이션으로부터 데이터를 시간별로 측정(또는 수집)하므로 이들 과학 데이터는 시간에 종속된 데이터이다. 많은 과학 응용의 시간 지원 데이터는 과학 분야의 특성에 따라 매우 정밀한 시간 단위로 수집하기도 하고 실험 시작부터 경과된 상대적인 시간에 따라 데이터를 수집한다. 달력상의 시간을 사용하는 기존의 시간 지원 데이터베이스는 과학 응용의 이러한 특징을 지원하지 못한다. 따라서 이 논문에서는 과학 실험 응용에서 요구하는 시간에 대한 특징과 요구 사항을 분석하고, 이들 과학 응용의 요구 사항을 만족하도록 확장한 새로운 시간 지원 데이터 모델을 제안한다. 이 모델에서는 실험이 경과된 상대적인 시간에 대한 데이터의 이력을 확장형 집합 개념을 사용하여 표현한다. 기존의 관계 데이터 모델과 유사하게, 이 모델은 집합 개념에 바탕을 두고 있으므로 데이터 모델과 그 연산의 의미를 쉽게 이해하고 사용할 수 있다는 장점이 있다.

---

+ 이 논문은 1992년도 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

\* 강원대학교 전자계산학과

\*\* 동의공업전문대학 전자계산과

## 1. 서론

항공 우주, 기상학, 천체학, 물리학, 생물학, 화학 등등의 과학 응용 분야에서는 실험 장비 또는 계측 기기를 통하여 해당 실험이나 관측에서 생성한 데이터를 수집하고 그 추세를 감시하며 해당 분야의 과학적 분석에 활용하고 있다. 이들 과학 응용에서 생성하는 데이터를 관리하는 데이터베이스를 과학 데이터베이스(scientific database)라고 한다 [Jones & Pfaltz 1990]. 과학 데이터베이스 관리에 필요한 요구사항들은 1990년 과학 데이터베이스 관리를 NSF 워크숍 [Jones & Pfaltz 1990]에서 깊이 논의된 후 이에 대한 관심과 연구가 점점 증가하고 있다 [Cushing et al. 1992, Cushing & Maier & Rao 1992, Hachem & Gennert & Ward 1993, Stonebraker et al. 1993].

이들 과학 응용 자료들은 실제계에 대한 관찰이나 물리적인 실험 장치 혹은 작동중인 기계장치나 시뮬레이션 장치 등으로부터 얻어지기 때문에 각 자료는 그것이 측정된 (또는 생성된) 시간과 연관된다 [Dreyer & Dittrich & Schmidt 1995, Jones & Pfaltz 1990, Segev & Shoshani 1987].

따라서 시간에 따라 변화한 데이터의 이력을 관리하는 것이 과학 데이터베이스의 가장 중요한 요구사항중의 하나이다. 이와 같은 이력 데이터 또는 시간 지원 데이터를 다루는 방법들은 이미 많이 연구가 진행되어 [McKenzie 1986, Segev & Jensen & Snodgrass 1995, Soo 1991, Tansel 1993], 여러 데이터 모델과 질의 언어가 개발되었으며 [Clifford & Tansel 1985, Clifford & Warren 1983, Gadia 1988, Jensen & Soo & Snodgrass 1994, McKenzie & Snodgrass 1991, Navathe & Ahmed 1989, Snodgrass 1987, Snodgrass et al. 1994, Tansel 1986], 여러 가지 저장 구조와 색인 기법 등이 개발되었고 [Ahn & Snodgrass 1988, Kokkotos & Syropoulos 1994, Lum et al. 1984, Rotem & Segev 1987], 몇몇 실제 프로토타입 시스템들이 개발되었다

[Segev & Jensen & Snodgrass 1995, Snodgrass 1987]. 이들 시간지원 데이터베이스는 데이터의 이력을 관리하기 위해 유효시간(valid time)과 거래시간(transaction time)의 두 종류 시간(또는 그 중의 한가지 시간)을 제공한다 [Jensen et al. 1992, Snodgrass & Ahn 1985]. 여기서 유효 시간은 데이터가 실제계에서 유효한 시간을 나타내고 거래시간은 그 데이터가 데이터베이스에 저장된 시간을 나타낸다.

과학 응용에서는 수 백만분의 일초(microsecond) 등의 매우 정밀한 단위로 데이터를 측정하기도 하고, 데이터가 발생한 달력상의 절대적인 시간 보다는 실험 시작부터 경과한 상대적인 시간 (또는 시간 순서(temporal order))이 더 중요한 경우가 많이 있다 [Segev & Shoshani 1987]. 그러나 기존의 시간지원 데이터베이스에서 사용하는 거래시간과 유효시간은 년/월/일 시:분:초의 달력상의 절대적인 시간을 사용하고 있어 과학 응용의 이러한 요구사항을 직접적으로 반영하지 못하고 있다. 이 논문에서는 과학 응용의 요구사항을 분석하고 이 요구사항을 반영할 수 있는 새로운 데이터 모델을 제안하였다. 과학 데이터의 요구사항은 화학 실험에서 많이 사용하는 가스 크로마토그래피에서 생성하는 데이터를 대상으로 분석하였다 [김강진 외 3명 1992, Christian & O'Reilly 1986]. 가스 크로마토그래피에서 생성하는 데이터는 실험 시작부터 경과된 상대적인 시간 (또는 시간 순서)에 따라 표현된다. 이 논문에서 이들 상대적인 시간 순서에 따른 데이터의 이력을 확장형 집합 개념[Childs 1977]을 사용하여 표현하는 새로운 데이터 모델을 제안하였다.

또 이 논문에서는 제안된 데이터 모델을 바탕으로 과학 데이터를 다루는 연산과 그 의미를 규명하였다. 집합 개념에 바탕을 둔 기존의 관계 데이터 모델과 마찬가지로, 이 데이터 모델은 확장형 집합이라고 하는 집합 개념을 사용하므로 과학 실험 데이터의 이력을 편리하게 표현할 수 있으며 데이터 모델과 연산의 의미를 쉽게 이해

하고 사용할 수 있다는 장점이 있다.

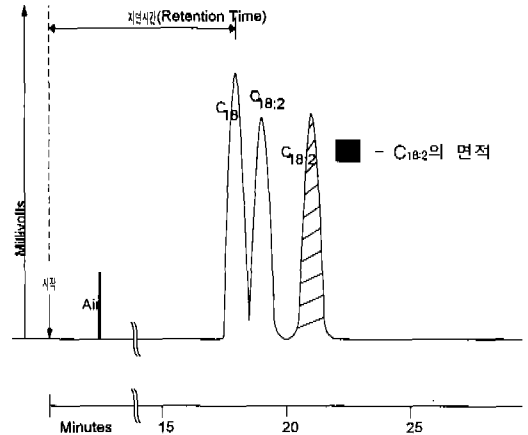
이 논문은 다음과 같이 구성되어 있다. 먼저 2장에서는 과학 실험 응용의 요구 사항을 분석하기 위해 가스 크로마토그래피를 사용한 분석 화학 실험의 특성에 대해 소개한다. 3장에서는 기존의 시간지원 데이터베이스의 특성과 그 문제점에 대해 살펴보고 이 문제점을 해결하기 위해 이 논문에서 제안하는 기본적인 방법에 대해 설명한다. 4장에서는 이 논문에서 제안하는 데이터 모델에 대한 정의를 소개하며, 5장에서는 이 데이터 모델을 바탕으로 시간 대수 연산을 설계한다. 끝으로, 6장에서는 연구 결과를 정리하고 결론을 맺는다.

## 2. 과학 실험 데이터 예제

이 장에서는 과학 실험에서 발생하는 시간 지원 데이터의 특성을 설명하기 위해 분석 화학 실험에서 많이 사용되고 있는 크로마토그래피에 대해 간단히 살펴보기로 한다. 크로마토그래피는 여러 가지 성분으로 구성된 혼합물의 성분(즉, 정성 분석)과 양(즉, 정량 분석)을 분석하는데 사용되는 화학 실험 장비로서 아미노산 및 단백질의 분리, 화학 공업에서의 분석 및 품질 관리 업무는 물론, 의약품 및 생물 화학 물질의 분석 및 대사 연구, 환경 오염 물질과 정도 분석등 매우 다양한 분야에 활용되고 있다 [김강진 외 3명 1992, Christian & O'Reilly 1986].

크로마토그래피의 기본적인 동작 원리는 다음과 같다. 분리할 혼합물 시료를 주입부(injection port)를 통해 주입하면 그 화합물을 구성하는 시료 성분 종류에 따라 일정 시간 지연(retain)시킨 후에 용출한다. 이때 시료 성분들이 용출된 시간과 양을 기록계에 기록하게 된다. 이렇게 기록된 그래프를 크로마토그램이라고 한다. 다음 <그림 1>은 크로마토그램의 예를 보여주고 있다. 동일한 실험 조건을 사용한다면 특정 시료 성분은 항상

동일한 지연 시간에 용출되며 그 양이 동일할 경우 같은 크기(즉 면적)를 갖는 봉우리가 생성된다



<그림 1> 크로마토그램 예제

크로마토그래피의 이러한 원리를 이용하여 미리 각 물질(성분)들에 대한 지연 시간과 표준 양에 대한 봉우리 면적들을 구해 데이터베이스로 구축해 둔 다음, 실험을 통해 수집된 특정화합물(예: 오염 물질, 의료 진단 샘플 등등)에 대한 크로마토그램을 얻어서 그 화합물을 구성하는 각 성분에 대한 지연시간과 봉우리 면적을 구한다. 이 지연 시간들과 동일한 지연 시간을 갖는 물질(성분)을 데이터베이스에서 검색함으로써 특정 시료(샘플)에 포함된 성분을 분석할 수 있다. 또 시료 화합물의 성분에 대한 봉우리 면적을 데이터베이스내의 저장된 면적과 비교하면 그 성분의 양도 측정할 수 있다.

크로마토그래피의 이러한 특성을 이용하면, 물이나 공기 시료를 추출하여 어떤 종류의 오염 물질이 기준치 이상의 양이 포함되었는지 여부를 검사하는데 유용하게 사용할 수 있다. 또 의약품 제조나 화학 공정의 품질을 감시하기 위해 생산품에 대한 샘플을 주기적으로 채취하여 각

성분이 정확한 비율로 포함되었는지 검사하는데 사용할 수도 있다. 이러한 성분 분석을 얼마나 정확하게 할 수 있는나 하는 것은 데이터베이스에 각종 물질에 대한 실험 데이터가 얼마나 풍부하고 정확하게 유지되어 있는나에 달려있다.

이 연구에서는 이러한 실험 데이터를 저장하고 관리하는 데이터베이스 관리 기술에 관해 연구하고자 한다. 이 데이터베이스가 관리하는 실험 데이터는 기본적으로 시간에 따라 측정(수집)되는 이력 데이터(또는 시간 지원 데이터)이다.

이 논문에서는 이러한 크로마토그래피 실험 데이터의 특성을 분석하고 이에 적합한 이력 데이터 관리 기법을 제안하고자 한다.

### 3. 실험 데이터의 요구 사항과 시간 지원 데이터베이스

#### 3.1 시간 지원 데이터베이스의 특징과 문제점

크로마토그래피에서 출력되는 데이터는 그 발생한 시간에 따라 순서대로 생성되며 유지된다.

이와 같이 시간에 따라 변화하는 데이터를 이력 데이터(historical data) 또는 시간 지원 데이터(temporal data)라고 한다 [Snodgrass & Ahn 1985]. 시간 지원 데이터는 화학 실험 등과 같은 각종 과학 실험뿐만 아니라 공장 자동화, 우주 탐사, 통신망 관리 등 각종 장치나 측정 기계의 상태를 시간별로 측정하는 응용에서 많이 발생한다. 이러한 시간 지원 데이터를 관리하는 데이터베이스를 시간지원 데이터베이스(temporal databases)라고 한다. 이 시간지원 데이터베이스에 대한 기술은 그 동안 많은 연구가 진행되어 여러 데이터 모델과 언어 및 프로토타입 시스템이 제안되었다 [Tansel et al. 1993, McKenzie 1986, Soo 1991, Segev & Jensen & Snodgrass 1995].

Snodgrass & Ahn[1985]은 시간 지원 데이터베이스에서 제공해야 할 시간 개념을 유효 시간(valid time)과 거래 시간(transaction time)의 두

가지로 소개한다. 여기서 유효 시간은 데이터(또는 사실)가 실제 세계에서 유효한 시간(즉, 발생한 시간)을 말하며, 거래 시간은 그 데이터(또는 사실)가 데이터베이스에 저장된 물리적인 시간을 말한다. 거래 시간은 데이터(또는 사실)를 데이터베이스에 저장한 시간을 이용하는 응용을 위해 사용되며, 한 데이터와 그 유효 시간이 여러 번 데이터베이스에 기록될 경우 그 변경된 내역(이력)을 거래 시간에 따라 유지할 수 있다. 따라서 거래 시간과 유효 시간은 서로 직교적(orthogonal)이다. 이 두 가지 시간은 이력 데이터를 관리하는 데이터베이스의 종류를 구분하는 기준이 된다[Jensen et al. 1992];

- (1) 스냅샷 릴레이션(snapshot relation) 또는 정적 릴레이션(static relation) : 기존의 릴레이션과 마찬가지로 아무런 시간 개념도 제공하지 않고 특정 시점의 상태만 유지하는 릴레이션.
- (2) 거래시간(transaction-time) 릴레이션 또는 롤백(rollback) 릴레이션 : 거래 시간에 따라 객체의 변화된 내역(이력)을 유지하는 릴레이션.
- (3) 유효시간(valid-time) 릴레이션 또는 이력(historical) 릴레이션 : 유효 시간에 따라 객체가 변화된 실제 세계의 내역(이력)을 유지하는 릴레이션.
- (4) 이원시간(bi-temporal) 릴레이션 또는 시간 지원(temporal) 릴레이션 : 유효 시간과 거래 시간 두 종류를 함께 제공하는 릴레이션.

위의 분류중에서 이원시간 릴레이션은 두 가지 시간 개념을 모두 지원하고 있어 실제 세계의 이력 데이터를 가장 충실하고 풍부하게 다룰 수 있다. 과학 실험 응용의 경우에도 실험 데이터가 실험이 진행되는(반응하는) 시간에 따라 발생하며, 동일한 실험을 여러 번 시행할 경우 동일한 객체(시료)에 대한 여러 번의 실험 데이터를 얻을 수 있다. 따라서 크로마토그래피 실험 데이터

를 관리하기 위해서는 시료의 반응 시간과 그 실험이 시행된 시간의 두 가지 직교적인 시간 개념을 필요로 한다. 이때 실험 시행 시간은 해당 실험이 실제로 언제 이루어진 시간을 나타내므로 시간지원 데이터베이스의 거래시간에 대응될 수 있으며, 시료의 반응시간은 각 시료 성분이 크로마토그램상에서 갖는 의미(효과)를 나타내므로 유효시간에 대응되는 개념이다. 따라서 이 논문에서는 과학 실험 데이터를 이원시간 릴레이션을 사용하여 관리하는 방법을 제안한다.

이때 과학 실험의 시행 시간은 이원시간 릴레이션의 거래 시간을 사용하여 모델링하며 시료의 반응 시간은 유효 시간을 사용하여 모델링하도록 한다.<sup>1)</sup> 그러나 기존의 이원시간 릴레이션에서 사용하는 시간은 과학 응용을 다루는데 필요한 다음과 같은 특성을 반영하지 못하고 있어 이들 특성을 수용하기 위한 확장이 필요하다.

첫째, 기존의 이원시간 릴레이션의 유효 시간은 년/월/일 시:분:초의 달력상의 시간(즉, 일상 생활에서 쓰고 있는 시간)을 다루고 있는데 반해, 실험 데이터는 실험이 시작된 시점에서부터 얼마의 시간이 경과했는가 하는 상대적인 시간이 필요하다. 즉 크로마토그램에서 한 물질에 대한 피크 값이 언제 발생했는가 하는 달력상의 시간은 아무런 의미가 없으며 실험 시작 시간으로부터 경과된 시간, 즉 상대적인 시간 순서가 중요하다. 피크 값이 발생한 달력상의 시간은 달라도 실험 시작에서부터 경과된 시간이 같으면

서로 동일한 물질이 되기 때문이다.

둘째, 몇몇 과학 실험 분야는 우리가 일반적으로 사용하는 시간 단위보다 훨씬 더 정밀한 시간 단위를 필요로 한다. 크로마토그래피 실험 데이터의 경우, 성분들의 피크 값을 정밀하게 나타내면 낼수록 물질(성분)들에 대한 변별력이 높아진다. 따라서 이력 데이터베이스에서 이러한 과학 실험 데이터를 다루기 위해서는 매우 정밀한 시간을 다룰 수 있도록 확장해야 한다.

셋째, 과학 실험은 보통 지속 시간이 몇 시간 혹은 며칠 정도의 시간이 소요되는 것이 보통이다. 예를 들면, 크로마토그래피에서는 실험에 소요되는 시간이 보통 30-40분 정도이며 길어도 수 시간 내에 실험이 종료된다. 이러한 짧은 기간내의 시간을 나타내는데 년/월/일 시:분:초의 달력상의 시간을 사용한다면 많은 시간 요소(즉, 년/월/일 등)가 동일한 값만 갖게 되고 따라서 사용하기 번거로울 것이다.

넷째, 실세계에서 사용하는 이력 데이터의 질의 유형은 크게 시점 질의(snapshot query)와 시간 순서 질의(temporal order query)의 두 가지 형태가 있다. 시점 질의는 1997년 8월 7일에 측정(수집)된 데이터는 무엇인가?와 같이 특정 시점의 값에 대한 질의이고, 시간 순서 질의는 50번째 발생한 데이터는 무엇인가?와 같이 시간 순서상의 값에 대한 질의이다. 이원시간 릴레이션의 시간들은 시점 질의 유형은 잘 처리할 수 있지만 시간 순서 질의는 직접적으로 처리하기 곤란하다. 과학 실험에서는 데이터의 발생 시간 뿐만 아니라 발생한 순서가 중요한 경우가 많아 시점 질의뿐만 아니라 시간 순서 질의도 효과적으로 제공할 방법이 필요하다.

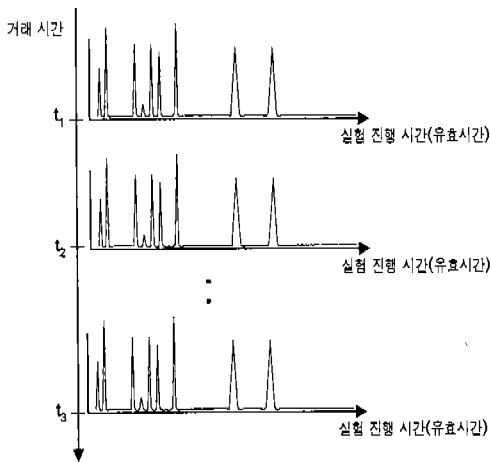
### 3.2 접근 방법

크로마토그래피 실험 데이터(즉, 크로마토그램)는 실험이 진행되는 시간에 따라 발생하는 이력 데이터이다. 또한 동일한 실험이 여러 번에

1) 일반적으로 거래시간은 데이터가 데이터베이스에 물리적으로 기록된 시간으로 정의된다. 그러나 이 논문에서는 실험이 시행된 시간은 변경될 수 없는 물리적인 의미가 있을 뿐만 아니라, 센서를 통해 컴퓨터에서 직접 실험 데이터를 수집하고 저장할 경우 실험 시행 시간과 거래 시간은 동일하므로 실험의 시행 시간을 거래 시간으로 모델링한다. 실험 시행과 데이터베이스 저장이 별도로 이루어질 경우 실험 시행 시간을 또 다른 유효 시간의 하나로 간주하여도 무방하다.

걸쳐 수행될 경우, 실험 시작(또는 실험 기간)만 다른 동일한 크로마토그램이 여러 번 생성된다.

따라서 크로마토그래피 실험 데이터를 저장/관리하기 위해서는 크로마토그램을 나타낼 시간 축외에도 그 실험이 언제 이루어졌는지 구분할 필요가 있다. 따라서 이 논문에서는 크로마토그래피 실험 데이터는 이원시간 릴레이션을 사용하여 유지한다. 즉, 이원시간 릴레이션의 거래 시간은 실험이 물리적으로 시작된(진행된) 시간을 표현하는데 사용하며, 유효 시간은 각 실험이 시작된 이후부터 실험이 경과된 시간을 나타내도록 한다.



<그림 2> 이원 시간에 의한 크로마토그래피 실험 데이터 표현

거래 시간은 실험이 개시된 시간(또는 기간)을 나타내며 보통 이 시간은 달력상의 시간의 의미로 주로 사용된다. 따라서 이원시간 릴레이션의 거래 시간을 실험 개시 시간을 나타내는데 그대로 사용할 수 있다. 유효 시간은 크로마토그래피에서 생성된 데이터의 생성 시간을 나타내는 것이다. 이 유효 시간은 실제 실험 진행(내용)과 관련된 시간이므로 달력상의 시간을 사용

하지 않고, 앞의 3.1절에서 설명한 바와 같이, 크로마토그래피 실험을 표현할 수 있도록 실험이 경과한 상대적인 시간을 나타내도록 한다.

이 상대적인 시간은 자연수 값을 사용하여 표현한다. 이 자연수 값은 실험 시작부터 경과한 시간을 나타내며 동시에 데이터가 발생한 순서를 나타낼 수 있다. 또 이 자연수를 실제 실험에 사용한 정밀한 시간 단위에 대응시켜 매우 정밀한 시간을 나타내는데도 사용할 수 있으며, 간단한 정수 값으로 시간을 표현하기 때문에 시간에 대한 연산이나 사용이 편리하다는 장점이 있다. 이 방법에서는 실험 데이터의 이력을 자연수 값에 대응하여 표현하므로 이를 확장형 집합(extended set) 개념[Chilids 1977]을 이용하여 이력 데이터를 정의하도록 한다.

일반적인 집합은 원소들에 대한 위치(또는 순서)를 다루지 않으므로 순서를 갖는 데이터 집합을 표현할 수 없다. 집합에 그러한 순서를 지원하기 위해 확장형 집합이 Chilids[1977]에 의해 제안되었다. 다음 정의에서 보는 바와 같이 확장형 집합은 집합의 원소에 순서 표시자를 대응시키는 집합 개념이다. 이 논문에서는 이 순서 표시자를 이용하여 크로마토그램의 유효 시간을 나타내는데 사용하도록 한다.

**【정의 1】** : 확장형 집합  $E$ 는 다음과 같이 정의된다 :  $E = \{a_1^{i_1}, a_2^{i_2}, \dots, a_n^{i_n}\}$ .

여기서  $a_j$  ( $1 \leq j \leq n$ )는 원소 식별자이고  $i_k$  ( $1 \leq k \leq n$ )는 순서 표시자이다. 순서 표시자  $i_k$ 는 자연수 값 중의 하나를 갖는다고 가정한다.

확장형 집합과 원소들의 관계는  $i$ -멤버십 관계에 의해 표현될 수 있다. 원소  $a$ 가 확장형 집합  $E$ 의  $i$ 번째 원소라는 것은 집합  $E$ 의  $i$ 번째 위치에  $a$ 가 있다는 의미이고 이를  $a \in_i E$ , 또는  $a^i \in E$ 로 표기한다.

예를 들면,  $E = \{a^1, b^2, c^3\}$ 은 확장형 집합이며 원소  $a$ 는 위치 1에,  $b$ 는 위치 2에,  $c$ 는 위치 3에 있

는 원소이다. 이는  $a \in_1 E$  (또는  $a^1 \in E$ ),  $b \in_2 E$  (또는  $b^2 \in E$ ),  $c \in_3 E$  (또는  $c^3 \in E$ ) 와 같이 표기될 수 있다.

이 논문에서는 시간에 따라 수집된 과학 실험 데이터를 확장형 집합을 사용하여 모델링한다. 확장형 집합의 순서 표시자를 실험 데이터가 발생한 순서 (또는 상대적 시간)를 나타내는데 사용한다. 즉 실험 시작으로부터  $i$  시간이 경과한 후의 값 (즉,  $i$ 번째 생성된 값)을 확장형 집합의  $i$ 번째 원소로 표현한다. 확장형 집합이라는 집합 개념을 사용하여 시간 순서를 표현하기 때문에 기존의 집합 개념에 익숙한 사용자들이 쉽게 이력 데이터를 이해하고 다룰 수 있는 장점이 있다. 또한 기존의 관계 데이터 모델이 집합에 바탕을 두고 있어 이 관계 데이터 모델을 약간만 수정함으로써 시간 순서를 갖는 이력 데이터를 위한 데이터 모델을 쉽게 정의하고 사용할 수 있다는 장점이 있다. 또 확장형 집합에서는 시간적 순서 (또는 실험 시작으로부터 상대적 경과 시간)를 표현할 수 있어 이러한 시간적 순서를 필요로 하는 과학 실험 데이터 표현이 가능하다.

#### 4. 데이터 모델 및 기본 정의

이 논문에서 제안하는 데이터 모델은 거래 시간과 유효 시간의 두 가지를 제공하는데, 거래 시간은 실험이 시행된 시간을 나타내는데 사용하며 달력상의 시간으로 표현한다. 유효 시간은 그 실험이 시작 시점부터 경과한 시간을 나타내는데 사용하며 확장형 집합의 순서 표시자를 이용하여 표현한다. 만약 동일한 실험이 여러 번 시행될 경우 각 실험 데이터는 동일한 유효 시간에 대해 데이터가 발생하며, 실험 시행 시간 즉 거래 시간만 다르다. 크로마토그램에서 거래 시간이 서로 달라도 피크 값에 대한 유효 시간이 같으면 같은 물질로 판정한다. 이 절에서는 이 데이터 모델에 대한 기본적인 기능과 정의에

대해 소개한다.

과학 실험 데이터를 표현하는 이원시간 데이터베이스의 거래 시간은 다음과 같은 전체 시간 영역 상에서 정의된다고 가정한다.

**【정의 2】** : 전체 시간 영역(universal time domain)  $T$ 는 일정한 간격의 시점들의 집합으로 정의된다.

즉  $T = \{t_0, t_1, \dots, t_i, \dots\}$ . 이때 시점의 간격은  $T$ 가 정의된 시간 단위(time unit)의 기본 시간이다. 전체 시간 영역  $T$ 는 선형 순서 집합(linear ordered set)이다. 따라서 만약  $i < j$ 이면  $t_i < t_j$ 이고 이것은  $t_i$ 가  $t_j$ 보다 더 이전 시점인 것을 의미한다.

이원시간 릴레이션의 두 가지 시간중 거래 시간은 반드시 이 전체 시간 영역,  $T$ , 내에 있는 시점들만 사용할 수 있다. 한 시간 주기  $[t_i, t_j]$ 는  $T$ 내에 있는 연속된 시점들의 집합으로 정의될 수 있다. 즉,  $[t_i, t_j] = \{t_k \mid t_i \in T, t_j \in T, t_k \in T, t_i \leq t_k \leq t_j\}$ . 또 사용자가 시간을 표현할 때 사용하는 단위는 년, 월, 일, 시, 분, 초등이 있는데 상위 단위로 표시된 시점은 하위 단위의 모든 시점을 포함하는 시간 주기를 나타낸다.

**【정의 3】** : 실세계에 존재할 수 있는 모든 가능한 객체들의 집합을  $O$ 라고 정의하자. 즉,  $O = \{o_1, o_2, \dots, o_m\}$ . 한 데이터베이스 DB란 이  $O$ 에 대한 표현(description)이다. 이때 한 DB내에는 여러 개의 릴레이션이 존재할 수 있다. 즉,  $DB = \{r_1, r_2, \dots, r_n\}$ . 여기서  $O_{r_i}$ 와  $O_{r_j}$ 를 각각 릴레이션  $r_i$ 와  $r_j$ 가 표현하는 객체들의 집합이라고 가정하자. 그러면  $O_{r_i} \cap O_{r_j} \neq \emptyset$  이며, 이 사실은  $O$ 내의 한 객체에 대한 정보가 여러 릴레이션에 나누어져 있음을 나타낸다.

**【정의 4】** : DB내의 각 릴레이션의 스키마( $R$ )는 애트리뷰트들의 집합으로 정의된다,  $R = \{A_1, A_2, \dots, A_n\}$ . 여기서  $\text{dom}(A_i)$ , 즉  $D_i$ 를 애트리뷰트

$A_i$ 의 영역(domain)이라고 하며, 이력 데이터를 표현할 수 있도록 확장형 집합을 사용한다. 애트리뷰트  $A_i$ 는  $O_r$ 과  $T$ 에서  $\text{dom}(A_i)$ 로의 함수로 정의된다. 즉,  $A_i : O_r \times T \rightarrow \text{dom}(A_i)$ .

한 애트리뷰트  $A_i$ 는 확장형 집합  $D_i (= \text{dom}(A_i))$ 내의 값들을 가진다. 애트리뷰트  $A_i$ 에 대해 유효 시간에 따라 변화한 이력 데이터는  $A_i$ 의 영역,  $D_i$ 의 값들의 연속으로 표현한다.

릴레이션 스킴  $R$ 에 대해 정의된 한 이력 릴레이션( $r$ )은 객체 집합  $O_r$ 에 대한 표현(description)의 집합이다. 즉,  $r = \{d_1, d_2, \dots, d_n\}$ . 여기서  $d_i$ 는  $O_r$ 내의 한 객체에 대한 표현이다. 각 표현에 대한 객체를 식별하는 함수를  $obj$ 라 정의한다. 예를 들어 만약  $d_i$ 가 객체  $o_i$ 의 표현이면,  $o_i = obj(d_i)$ 이다.

**[정의 5]** : 각 집합  $D_i$  ( $1 \leq i \leq n$ )는 확장형 집합으로 릴레이션 스킴  $R$ 내의 각 애트리뷰트  $A_i$ 의 영역이라고 하면, 한 릴레이션  $r$ 은 교차곱

$O_r \times \prod_{i=1}^n D_i \times T$ 의 부분 집합으로 정의된다. 즉,

$$r = \{ \langle o_i, v_1^{j_1}, v_2^{j_2}, \dots, v_n^{j_n}, t \rangle \mid o_i \in O_r, t \in T, v_k^{j_k} \in D_k, D_k = \text{dom}(A_k), v_k^{j_k} = d_i[A_k](t), 1 \leq k \leq n \}.$$

여기서  $d_i[A_k](t)$ 는 객체  $o_i (= obj(d_i))$ 가 시간  $t$ 에서 갖는 애트리뷰트  $A_k$ 의 값이다.

위 정의 중에서 같은 객체 (예:  $o_i$ )에 대해 갖는 모든 순서쌍의 집합을 객체 표현 ( $d_i$ 로 표기)이라 부른다.

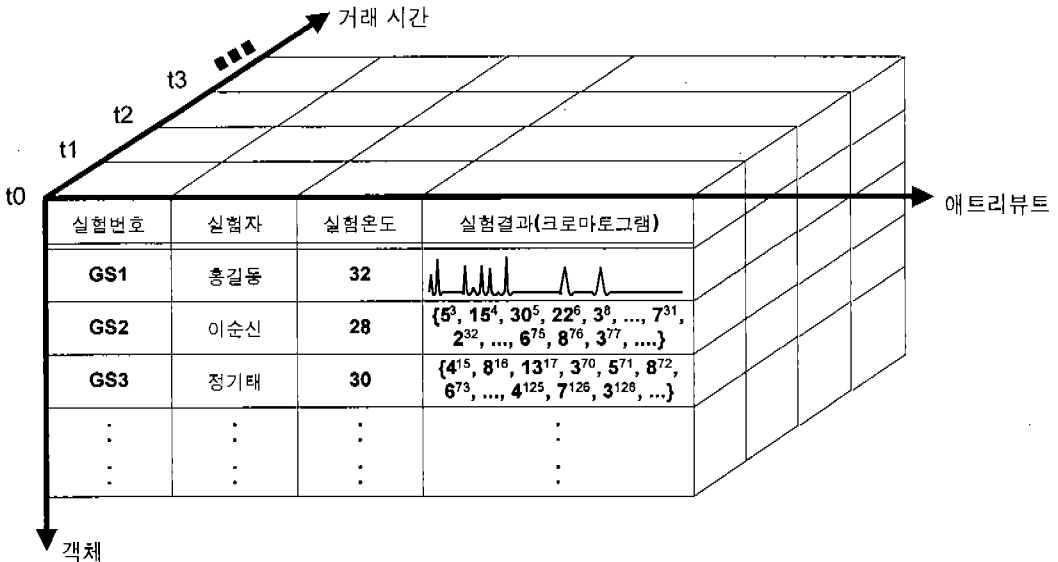
$$\text{즉 } d_i = \{ \langle o_i, v_1^{j_1}, v_2^{j_2}, \dots, v_n^{j_n}, t \rangle \mid t \in T, o_i \in O_r, v_k \in D_k, D_k = \text{dom}(A_k), o_i = obj(d_i), v_k = d_i[A_k](t), 1 \leq k \leq n \}.$$

한 객체 표현  $d_i$ 가 각 시점에서 갖는 값, 즉  $d_i(t)$ 을 거래시간 이력 어커런스(transaction time temporal occurrence)라고 부른다.  $d_i[A](t)$ 는 객체  $o_i$ 가 거래시간  $t$ 에서 갖는 애트리뷰트  $A$ 의 유효 시간 이력을 나타낸다.

$$\text{즉, } d_i[A](t) = \{ \langle o_i, v^j \rangle \mid o_i \in O_r, j \in N, v^j \in D, D = \text{dom}(A), o_i = obj(d_i) \}.$$

여기서,  $N$ 은 자연수 집합을 의미한다.

위에서는 제안된 데이터 모델에 대한 정의를



<그림 3> 이원시간 릴레이션 화학실험의 입방체 표현



설명하였다. 이 모델은 각 객체의 거래 시간 이력을 전체 시간 영역 T상의 시점을 사용하여 표현한다. 따라서 객체들이 거래 시간축에 따라 3차원 입방체 모양을 갖는 것으로 간주할 수 있다 (위의 <그림 3> 참조). 이 입방체의 한 셀 (cell)은  $d_i[A_i](t)$ 로 표기되며, 거래 시점 t에서 수집된 애트리뷰트 A의 유효시간 이력을 나타내는 집합(또는 리스트)이다. 따라서 이 모델의 이원시간 릴레이션에 대한 사용자 관점의 구조는 셀 값으로 원자 값이나 집합(또는 리스트)를 갖는 3차원 입방체 구조로 표현한다. 위의 <그림 3>은 이원시간 릴레이션 화학실험(실험번호, 실험자, 실험온도, 실험결과)에 대한 구조를 보여주는 예이다.

## 5. 시간 대수 연산(Temporal Algebra)

앞장에서는 데이터 모델에 대한 정의와 기호에 대해 설명하였다. 이 장에서는 이러한 정의와 기호를 이용하여 이 모델에 대한 시간 대수 연산들을 정의한다. 이 시간 대수 연산자들은 세 가지 종류로 구분된다: (1) 일반 대수 연산, (2) 거래 시간 대수 연산, (3) 시간 순서 대수 연산. 이 장에서 사용하는 모든 연산은 그림 3에 나와 있는 3차원 입방체를 사용하여 설명하며, 모든 예제 질의들은 릴레이션 화학실험(실험번호, 실험자, 실험온도, 실험결과)을 사용하여 작성한다.

### 5.1 일반 대수 연산

#### (1) 선택선(Selection)

특정한 조건을 만족하는 객체들에 대한 정보를 선택하는 연산이다. 즉, 객체 축 방향의 부분 집합을 구한다. 선택 조건은 대개  $A_i \theta v_i$  형태로 된 항들의 조합으로 구성된다. 여기서  $A_i$ 는 한 릴레이션(r)에 대한 스키마(R)에 속하는 한 애트리뷰트이며,  $v_i$ 는  $dom(A_i)$ 내에 속하는 한 값이다.

또  $\theta$ 는 비교 연산자  $=, \neq, <, \leq, \geq,$

$>$  중의 하나를 표시한다. 이 선택선 연산자( $\sigma$ )는 다음과 같이 정의된다.

**【정의 6】** :  $\sigma_p(r) = \{d_i \mid d_i \in r, t_j \in T, p(d_i, t_j)\}$ , 여기서 p는 선택 조건식을 나타내며  $p(x)$ 는 x가 조건식을 만족하면 참을 반환한다.

예 1 : '홍길동'이 실험한 모든 실험 내용을 검색하라.

$\sigma_{\text{실험자}='홍길동'}(\text{화학실험})$

#### (2) 프로젝션(Projection)

각 객체들에 대해 원하는 애트리뷰트만 결과로 선택하는 연산이다. 즉, 애트리뷰트 축 방향의 부분 집합을 구한다. 프로젝션 연산

$\pi_{A_i, A_{i+1}, \dots, A_j}(r)$ 은 각 객체의  $A_i, A_{i+1}, \dots, A_j$  애트리뷰트에 대한 값만 결과에 포함시킨다.

**【정의 7】** : X는 애트리뷰트들의 집합이며 릴레이션 스키마 R의 부분 집합이라고 하자. 프로젝션 연산자는 다음과 같이 정의된다:  $\pi_X = \{d_i \mid X \mid d_i \in r\}$ .

예 2 : 실험번호가 'GC5'인 실험에 대한 실험자와 실험결과를 검색하라.

$\pi_{\text{실험자}, \text{실험결과}}(\sigma_{\text{실험번호}='GC5'}(\text{화학실험}))$

#### (3) 조인(Join)

조인은 두 릴레이션을 결합하는 연산이다. 이때 조인 조건으로 사용하는 애트리뷰트는 동일한 영역을 가져야 한다. 조인 조건으로는 임의의 비교 연산자가 가능하며 이러한 조인을  $\theta$ -조인이라 한다.

**【정의 8】** : r은 릴레이션 스키마  $R_r = \{A_1, A_2, \dots, A_m\}$ 에 대한 릴레이션이고, s는 스키마  $R_s = \{B_1, B_2, \dots, B_n\}$ 에 대한 릴레이션이라고 하자. 조인 연산은 다음과 같이 정의된다:

$$r \times_{A, \theta B}, s = \{d_k \mid d_r \in r, d_s \in s, t_r, t_s \in T, \\ d_r[A_i](t_r) \theta d_s[B_j](t_s), (A_r \in R_r, B_s \in R_s, \\ d_k[A_r](t_r) = d_r[A_r](t_r), d_k[B_s](t_s) = d_s[B_s](t_s), \\ d_k[s.tt](t_r) = t_s)\}.$$

여기서  $s.tt$ 은 릴레이션  $s$ 의 각 객체 표현  $d_s$ 의 거래 이력 어커런스들에 대한 거래 시간을 유지하는 애트리뷰트이다.

예 3 : 실험번호가 'GC5'의 실험자의 이름, 소속, 주소를 검색하시오.

$\pi_{\text{실험자, 소속, 주소}}(\text{화학실험} \times \text{실험자} = \text{이름 연구원})$

#### (4) 집합 연산자

합집합(union), 차집합(difference), 교차곱(cartesian product)의 세 가지 집합 연산에 대한 정의를 설명한다. 나머지 일반 집합 연산은 이들 연산으로부터 유도할 수 있으며 비슷한 방법으로 정의할 수도 있다. 먼저 합집합( $\cup$ )은 두 릴레이션의 객체 표현(object description)을 합치는 연산이다. 다음 정의에서 보는 바와 같이, 합집합은 각 거래 시점에서 애트리뷰트 값(즉, 유효시간 이력)들을 합친다.

**[정의 9]** : 릴레이션  $r$ 과  $s$ 는 스키  $R=\{A_1, A_2, \dots, A_m\}$ 에 대한 릴레이션이라고 하자. 합집합 연산은 다음과 같이 정의된다 :

$$r \cup s = \{d_q(t) \mid t \in T, d_r \in r, d_s \in s, \\ A_i \in R, d_q[A_i](t) = d_r[A_i](t) \cup d_s[A_i](t), \text{obj}(d_r) = \text{obj}(d_s)\}.$$

차집합(-)은 한 릴레이션에는 존재하지만 다른 릴레이션에서는 존재하지 않는 객체 표현을 찾는 연산이다.

**[정의 10]** : 릴레이션  $r$ 과  $s$ 는 스키  $R=\{A_1, A_2, \dots, A_m\}$ 에 대한 릴레이션이라고 하자. 합집합 연산은 다

음과 같이 정의된다 :

$$r - s = \{d_q(t) \mid t \in T, d_r \in r, (s\text{내에} \\ \text{obj}(d_r) = \text{obj}(d_s)\text{인 } d_s\text{가 없다면, } d_q(t) = d_r(t) \\ \text{또는 } (s\text{내에 } \text{obj}(d_r) = \text{obj}(d_s)\text{인 } d_s\text{가 있다면,} \\ d_q[A_i](t) = d_r[A_i](t) - d_s[A_i](t), \\ A_i \in R)\}.$$

교차곱( $\times$ )은 릴레이션  $r$ 과  $s$ 에 있는 객체 표현의 거래이력 어커런스들의 모든 가능한 조합을 구하는 연산이다. 이 조합들은  $r$ 의 거래 시간을 기준으로 정렬하며  $s$ 의 거래시간은 별도의 애트리뷰트에 유지한다.

**[정의 11]** :

릴레이션  $r$ 은 스키  $A=\{A_1, A_2, \dots, A_m\}$ ,  $s$ 는 스키  $B=\{B_1, B_2, \dots, B_n\}$ 에 대한 릴레이션이라고 하자. 교차곱은 다음과 같이 정의된다 :

$$r \times s = \{d_q \mid d_r \in r, d_s \in s, t \in T, d_q[A](t) = \\ d_r[A](t), (t_s \in T, d_q[B](t) = d_s[B](t_s), \\ d_q[s.tt] = t_s)\}.$$

여기서  $s.tt$ 은 릴레이션  $s$ 의 각 객체 표현  $d_s$ 의 거래 이력 어커런스들에 대한 거래 시간을 유지하는 애트리뷰트이다.

## 5.2 거래 시간 대수 연산

### (1) 거래 이력 선택(selection)

거래 이력 선택( $\eta$ 로 표기함)은 원하는 거래 시간에 대한 정보만 추출하기 위해 거래 시간축 방향의 부분 집합을 구하는 연산이다. 즉, 원하는 거래 시간에 대한 거래 이력 어커런스들만 추출하는 연산이다.

거래시간  $t$ 에서 릴레이션  $r$ 의 상태는  $r(t)$ 로

표기하며,  $r(t) = \{d_i(t) | d_i \in r, t \in T\}$  이다. 여기서  $d_i(t)$ 는 각 객체 표현  $d_i$ 가 거래 시간  $t$ 에서 갖는 거래 이력 어커런스이다. 원하는 거래 시간에 대한 조건이 일정 기간(time interval)으로 주어지면 그 기간내의 모든 거래 이력 어커런스들이 결과에 포함된다.

**[정의 12]** :  $\eta_\nu(r)$ 를 시간 원소  $\nu$ 에 대한 릴레이션  $r$ 의 거래 이력 실택션이라고 하자. 그러면,  $\eta_\nu(r) = r(\nu) = \{d_i(t) | d_i \in r, t \in T, t \in \nu\}$ 이다. 여기서  $\nu$ 는 거래시점들의 집합을 의미한다.

예 4 : 1997년 9월 18일에 실시한 실험번호가 'GC5'인 실험에 대한 정보를 검색하라.

$\tau_{1997/9/18}(\sigma_{\text{실험번호}='GC5'}(\text{화학실험}))$

(2) 거래 시간 추출(transaction time projection)

이력 데이터의 거래 시간 자체도 유용한 정보 중의 하나이다. 각 객체들의 거래 이력들이 발생한 시간을 추출하는 연산을 거래 시간 추출 연산이라 하며 다음과 같이 정의된다:

**[정의 13]** :  $\tau$ 는 거래 시간 추출 연산을 나타내며 이 연산의 매개 값들의 종류에 따라 다음과 같은 결과를 반환한다:

(a)  $\tau(d_i(t)) = t$ , 즉 거래 이력 어커런스에 대한 거래 시간을 반환한다;

(b)  $\tau(d_i) = \{t | t \in T, d_i(t) \in d_i\}$ , 즉 한 객체 표현의 거래 이력이 존재하는 기간을 반환한다.

(c)  $\tau(r) = \bigcup_{d_i \in r} \tau(d_i)$ , 즉 릴레이션  $r$ 의 거래 이력이 존재하는 기간을 반환한다.

예 5 : 실험번호가 'GC5'인 실험이 실시된 시간(기간)은 언제인가?

$\tau(\sigma_{\text{실험번호}='GC5'}(\text{화학실험}))$

### 5.3 시간 순서 대수 연산

한 객체 표현( $d$ )이 각 거래 시점( $t$ )에서 갖는 애트리뷰트( $A$ )내에 유지되는 값(즉,  $d[A](t)$ )은 그 애트리뷰트에 대한 유효 시간 이력이다.

이 유효 시간 이력은 실험이 진행된 상대적인 경과 시간에 따라 데이터의 이력을 유지한 것이며 확장형 집합을 사용하여 이를 표현한다. 따라서 유효 시간 이력은 확장형 집합이라고 하는 집합 개념으로 간단히 표현되며, 기존의 집합 연산자들을 사용하여 유효 시간 이력을 다룰 수 있다.

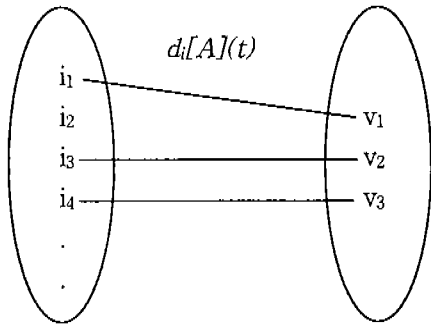
확장형 집합도 집합의 각 원소들이 위치 표시자(또는 시간 순서)를 갖는다는 점을 제외하고는 기존의 집합 개념과 동일하므로 이들 연산자들을 쉽게 확장하고 이해할 수 있을 것이다.

또 기존에 널리 사용되는 관계 데이터 모델과 유사하게 이 데이터 모델이 집합 개념을 사용하므로 사용자들이 쉽게 유효 시간 이력의 연산들을 이해하고 다룰 수 있다는 장점이 있다. 이 절에서는 이 유효 시간 이력을 다루는 연산에 대해 소개한다.

앞에서 설명했듯이,  $d_i[A](t)$ 은 객체  $o_i$ 가 거래 시간  $t$ 에서 애트리뷰트  $A$ 에 대해 가지고 있는 유효시간 이력을 나타내며, (여기서  $d_i$ 는 객체  $o_i$ 에 대한 표현이라고 가정한다) 다음과 같이 상대적 시간 순서와 값의 쌍으로 정의된다 :

$$d_i[A](t) = \{v^j | v^j \in \text{dom}(A), v \text{ is the value valid at } j\}.$$

이 정의에서 보는 바와 같이,  $d_i[A](t)$ 는 확장형 집합의 하나이며 상대적 시간 순서와 값들간의 사상으로 간주할 수 있다. 다음 그림 4는 이 사상에 대한 개념을 설명하고 있다.



$tdom(d_i[A](t))$                        $vdom(d_i[A])$

<그림 4>  $d[A](t)$ 에 대한 사상의 예

【정의 14】 :  $d_i$ 를 객체  $o_i$ 의 객체 표현이라고 하자. 그러면  $d_i[A](t)$ 는 유효 시간 이력을 나타내는 확장형 집합으로 표현되며, 확장형 집합에 대한 두 가지 함수  $tdom$ 과  $vdom$ 는 다음과 같이 정의된다:

$$tdom(d_i[A](t)) = \{j | v^j \in d_i[A](t)\},$$

$$vdom(d_i[A](t)) = \{v | v^j \in d_i[A](t)\}.$$

확장형 집합으로 표현되는 유효시간 이력의 각 원소들에 대한 시간 순서 값 자체를 반환하는 함수로는  $torder$ 를 사용한다. 즉,  $torder(v^j) = j$ 이다. 이 함수는 원소의 유효 시간을 찾기 위해 필요하며 실렉션 연산에서 이 유효 시간을 조건으로 정보를 검색하고자 할 때 사용할 수 있다. 이것은 거래시간 추출 연산에 대응하는 연산이다.

임의의 상대적 유효 시간(또는 시간 간격)에 대한 이력을 검색하는 연산으로 유효이력 실렉션( $\omega$ 로 표기함)을 사용할 수 있다. 이에 대한 의미는 정의 15에 나와 있다. 이 연산 외에도 상대적 유효 시간을 시간 순서(temporal order)로 취급하는 다음과 같은 시간 순서 연산자를 사용할 수도 있다 : *first*, *second*, *3rd*, ..., *n-th*, *current - 1*, *current*, *oldest* (또는 *smallest*),

*latest* (또는 *largest*). 이 중에서 *first* 함수는 상대적 유효 시간이 1인 원소를 반환하며, *n-th* 함수는 유효 시간이  $n$ 인 원소를 반환한다. 나머지 함수들도 이와 동일한 원리로 동작한다.

【정의 15】 :  $\omega_v (d_i[A](t))$ 를 유효 시간 원소  $v$ 에 대한 유효 이력 실렉션이라고 하자. 그러면,  $\omega_v (d_i[A](t)) = \{v^j | v^j \in d_i[A](t), j \in v\}$ 이다. 여기서  $v$ 는 유효 시간 값들의 집합을 의미한다.

예 6 :

(a) 1997년 9월 18일에 실시한 실험번호가 'GC5'인 실험이 실험 시작 후 100ms 경과한 시점의 값을 구하시오. (여기서 실험 데이터는 1ms 간격으로 수집된다고 가정한다.)

$$\omega_{100} (\pi_{\text{실험결과}} (\sigma_{\text{실험번호}='GC5'} (\eta_{1997/9/18} (\text{화학실험}))))$$

(b) 1997년 9월 18일에 실시한 실험중 실험결과 의 첫 번째 값이 0이 아닌 실험의 실험번호를 구하시오.

$$\pi_{\text{실험번호}} (\sigma_{\text{first}(\text{실험결과}) \neq 0} (\eta_{1997/9/18} (\text{화학실험}))))$$

(c) 1997년 9월 18일에 실시한 실험번호가 'GC5'인 실험의 최대(피크)값이 발생한 유효 시간을 구하시오.

$$torder(\max (\pi_{\text{실험결과}} (\sigma_{\text{실험번호}='GC5'} (\eta_{1997/9/18} (\text{화학실험}))))))$$

## 6. 결 론

분석 화학등 많은 과학 실험은 시간에 따라 데이터를 수집하므로 이력 데이터에 대한 관리가 필요하다. 이 논문에서는 기존의 이원시간 데이터

모델을 확장하여 과학 실험에서 발생하는 데이터를 위한 데이터 모델을 제안하였다. 기존의 이원시간 데이터 모델이 년/월/일 시:분:초의 달력상의 시간만을 사용하는데 반해, 실험 데이터는 달력상의 시간보다 실험 시작부터 얼마나 경과해서 발생했는가 하는 상대적 (경과) 시간이 더 중요하다. 이 모델에서는 이러한 실험 데이터가 발생한 상대적 시간 순서에 따라 이력을 표현할 수 있도록 하였다. 이 상대적 시간 순서는 확장형 집합 개념을 사용하여 모델링하였다.

또 이 모델에서는 상대적 시간 순서 외에도 실험이 시행된 시간 (즉, 달력상의 시간)을 거래 시간으로 표현한다. 이 거래 시간은 동일한 실험이 여러 번 반복될 경우, 이들을 서로 구분하기 위해 사용된다.

또한 이 논문에서는 제안된 모델을 바탕으로 과학 실험용 이력 데이터를 다루는데 필요한 연산을 제공하였다. 이 연산들은 (1) 일반적인 집합 연산을 확장한 것 외에도 (2) 거래 시간 이력을

다루는 연산과 (3) 상대적 시간 순서에 따라 표현되는 유효 시간 이력을 다루는 연산등으로 이루어진다. 기존의 관계 데이터 모델과 유사하게, 이 모델은 집합 개념에 바탕을 두고 있으므로 모델과 연산의 의미를 이해하고 사용하기가 쉽다는 장점이 있다.

이 논문에서는 과학 실험에서 생성되는 이력 데이터를 관리하는 데이터 모델에 대해서만 연구하였다. 앞으로 이 모델을 바탕으로 과학 데이터를 다루는 데이터베이스 언어를 설계하는 연구를 수행할 예정이다. 또, 이 논문에서 설계한 기본적인 연산자 외에도 다양한 과학 분석 기능을 위한 함수를 제공하고, 과학 분석을 쉽게 수행하도록 지원하는 기능들(예: 그래픽 사용자 인터페이스, 시뮬레이터, 과학 분석 모듈 등)에 대해서 연구하는 것이 필요하다. 나아가 이 모델에서 제공하는 과학 응용을 위한 시간 지원 데이터의 특성과 개념을 바탕으로 실제 시스템 개발에 필요한 기술들을 연구할 수 있을 것이다.

## 〈참고문헌〉

- 김강진외 3명, 「분석화학」, 자유 아카데미, pp.635-678, 1986.
- Ahn, I. and Snodgrass, R., "Partitioned Storage for Temporal Databases," *Information Systems*, Vol.13, No.4, pp.369-391, 1988.
- Childs, D. L., "Extended Set Theory: A General Model of Very Large, Distributed, Backend Information Systems," *Proc. of VLDB Conference*, pp.28-46, 1977.
- Christian, G. D. and O'Reilly, J. E., *Instrumental Analysis*, Allyn and Bacon., pp.658-765, 1986.
- Clifford, J. and Tansel, A. U., "On An Algebra for Historical Relational Databases: Two Views," *Proc. of ACM SIGMOD Conference*, pp.247-265, 1985.
- Clifford, J. and Warren, D. S., "Formal Semantics for Time in Databases," *ACM Trans. on Database Systems*, Vol.8, No.2, pp.214-254, 1983.
- Cushing, J. B. et al., "Object-Oriented Database Support for Computational Chemistry," *Proc. of Conference on Statistical and Scientific Database Management*, pp.58-76, 1992.
- Dreyer, W., Dittrich, A. K., and Schmidt, D., "Research Perspectives for Time Series Management Systems," *ACM SIGMOD Record* Vol.24, No.4, pp.10-15, 1995.
- French, J. C., Jones, A. K., and Pfaltz, J. L., "Summary of the Final Report of the NSF Workshop on Scientific Database Management," *ACM SIGMOD Record*, Vol.19, No.4, pp.32-40, 1990.

- Gadia, S. K., "A Homogeneous Relational Model and Query Language for Temporal Database," *ACM Trans. on Database Systems*, Vol.13, No.4, pp.418-448, 1988.
- Hachem, N. I., Gennert, M. A., and Ward, M. O., "An Overview of the Gaea Project," *IEEE Data Engineering Bulletin*, Vol.16, No.1, pp.9-13, 1993.
- Jensen, C., Soo, M., and Snodgrass, R., "Unifying Temporal Data Model via a Conceptual Model," *Information Systems*, Vol.19, No.7, pp.513-547, 1994.
- Jensen, C. S. et. al., "A Glossary of Temporal Database Concepts," *ACM SIGMOD Record*, Vol.21, No.3., 1992.
- Kokkotos, S. and Spyropoulos, C. D., "A Framework for Developing Temporal Databases," *Proc. of DEXA Conference, Lecture Notes in Computer Science 856*, Springer-Verlag, 236-245, 1994.
- Lum, V. et. al., "Designing DBMS Support for the Temporal Dimension," *Proc. of ACM SIGMOD 84*, pp.115-130, 1984.
- McKenzie, E., "Bibliography: Temporal Database," *ACM SIGMOD Record*, Vol.15, No.4, pp.40-52, 1986.
- McKenzie, E. and Snodgrass, R., "An Evaluation of Relational Algebras Incorporating the Time Dimension in Databases," *ACM Computing Surveys*, Vol.23, No.4, pp.501-543, 1991.
- Navathe, S. B. and Ahmed, R., "A Temporal Relational Model and A Query Language," *Information Sciences*, Vol.49, pp.147-175, 1989.
- Rotem, D. and Segev, A., "Physical Organization of Temporal Data," *Proc. of ICDE Conference*, pp.547-543, 1987.
- Segev, A., Jensen, C. S., and Snodgrass, R. T., "Report on The 1995 Int'l Workshop on Temporal Databases," *ACM SIGMOD Record*, Vol.24, No.4, pp.46-60, 1995.
- Segev, A. and Shoshani, A., "A Logical Modeling of Temporal Databases," *Proc. of ACM SIGMOD Conference*, 1987.
- Snodgrass, R. and Ahn, I., "A Taxonomy of Time in Databases," *Proc. of ACM SIGMOD Conference*, pp.236-246, 1985.
- Snodgrass, R. et al., "TSQL2 language specification," *ACM SIGMOD Record*, Vol.23, No.1, pp.65-86, 1994.
- Snodgrass R., "The Temporal Query Language Tquel," *ACM Trans. on Database Systems*, Vol.12, No.2, pp.247-298, 1987.
- Soo, M. D., "Bibliography on Temporal Databases," *ACM SIGMOD Record*, Vol.20, No.1, pp.14-23, 1991.
- Stonebraker, M., Chen, J., Nathan, N., Paxson, C., and Wu, J., "Tioga: Providing Data Management Support for Scientific Visualization Applications," *Proc. of VLDB Conference*, pp.25-38, 1993.
- Tansel, A. U., "Adding Time Dimension to Relational Model and Extending Relational Algebra," *Information Systems*, Vol.11, No.4, pp.343-355, 1986.
- Tansel A. et al., *Temporal Databases : Theory, Design, and Implementation*, The Benjamin/Cummings Publishing Company Inc., 1993.