

데이터베이스 설계(I)

-개념적 데이터베이스 설계 Conceptual Database Design



노봉남
전남대학교 전산학과 교수
Noh, bong-nam.
Dept. Computer Science
Chonnam National
University.

요 즘에는 정보 사회란 말을 너무나 자주 듣게 되어 식상하게 한다. 아무리 작은 업무에서도 컴퓨터를 이용한 데이터의 관리 및 결과물의 이용은 일상사처럼 되어 버렸다. 작은 업무라고 일컬어지는 것조차 수 십만 메가 바이트의 데이터를 유지하고 관리하는 데이터베이스 시스템이 존재한다. 그러나 시스템의 운영 상에 빈번한 과부하를 경험하게 된다. 증가하는 데이터를 관리하지 못하고 시스템의 속도가 느려지거나 급기야는 고장을 일으키는 불상사를 간혹 목격하게 된다. 왜 이럴까. 그것은 전산화 단계에 있어서 기존에 시스템을 분석하고 이를 바탕으로 데이터베이스 시스템을 구축하는 과정에서 임기 응변식의 분석과 설계 구현이 이루어졌기 때문이다. 데이터베이스 시스템 구축하기 위한 기존 시스템을 요구 사항 수집 및 분석하고, 데이터베이스 시스템을 계획·설계하는 과정을 거쳐 시스템의 구현하는 일련의 과정은 최대한 정확히 요구를 반영할

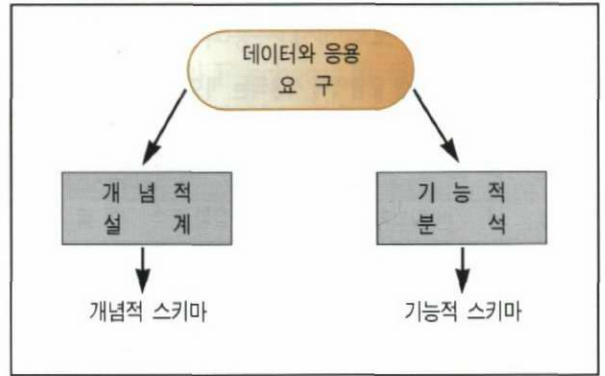
수 있어야 한다. 이러한 과정에서 데이터베이스의 설계는 데이터베이스의 구조와 내용, 그리고 수행될 응용을 포함하는 데이터베이스의 구성을 결정하는 과정이다. 이러한 설계 과정이 정확히 이루어졌을 때 운영상에 문제를 포함하지 않는 데이터베이스 시스템의 개발이 가능하다.

1. 데이터베이스 설계를 위한 모델

오래전에, 데이터베이스 설계는 경험에 의한 작업으로써 과학적이기 보다 경험적인 것으로 여겨졌다. 또한, 데이터베이스 시스템들이 상업적이면서도 과학적이고 기술적인 응용을 만족시키면서 점차 여러 분야에 사용되게 되었다. 이러한 응용의 확장은 데이터베이스 설계에 새로운 연구를 자극하게 되었다. 근래에는, 데이터베이스 설계에 대한 많은 연구가 이루어지면서 다양한 방법과 수단을 갖춘 기법들이 연구되기에 이르렀다. 데이

터베이스 설계에 사용되는 데이터 모델은 데이터의 추상화 정도에 따라 개념적 모델, 논리적 모델, 그리고 물리적 모델로 구분된다. 1970년대 말부터 몇몇 개념적 데이터베이스 모델에 대한 연구들이 발표되었다. 대표적으로 의미적 데이터 모델 (semantic data model), 구조적 모델(structural model), 기능적 모델(functional model) 등이 있으며 이러한 모델들은 추상화 기법(abstraction mechanism)에 기반을 두고 있다. 특히 개체 관련성 모델 (Entity Relationship model)은 개념적 데이터 표현을 위한 도구로 많은 부분에서 표준으로 사용되고 있다. 개체 관련성 모델은 단지 소수의 모델링 개념을 근간으로 하며 매우 효과적인 그래픽 표현을 가지고 있고 모델에서 각 요소들은 그래픽 기호들로 변환될 수 있다. 그러나 이러한 모델링 기법은 데이터 중심 접근(Data Driven Approach)으로서 데이터와 데이터의 특성에만 중점을 둔 접근 방법이다. 데이터 중심 접근 방법에 있어 먼저 데이터베이스가 설계된 후 데이터베이스를 사용하는 응용이 설계된다. 이러한 기법은 1970년대 말에 데이터베이스 기법으로 개발되었다. 또다른 시스템 설계를 위한 접근 방법으로서 데이터보다 응용에 중점을 둔 설계 방법으로서 SADT Diagrams, ISAC Activity Graphics(A-Graphs), Petri Nets, Information Control Nets(ICNS) 등이 있으며 가장 널리 사용되는 설계 방법으로는 1960년대 개발된 데이터 흐름 모델(Data Flow Model)이 있다. 이는 기능 중심 접근(Functional Driven Approach)으로서 정보의 흐름을 주로 표현한다. 데이터 중심 접근 방법과 기능 중심 접근 방법은 시스템을 설계하는데 상호 보완적이다. 즉 이들은 서로에게 훌륭한 특성을 제공할 수 있으며 긴밀하게 관련됨으로써 정확한 시스템 설계를 가능하게 한다. 이러한 면에서 데이터베이스를 설계하기 위한 방법은 그림 1과 같이 데이터 중심 접근 방법과 기능 중

(그림 1) 시스템 설계를 위한 데이터 중심 접근과 기능 중심 접근의 조합



심 접근 방법을 조합한 방법으로, 두개의 설계 방법은 서로에게 영향을 반영하도록 함으로써 상호 일관성 있고 완벽한 데이터베이스 설계가 가능하다.

많은 데이터베이스 설계자들에게 익숙한 개념적 설계 기법들로 먼저 데이터 중심 접근 방법으로 대표적인 개체 관련성 모델의 개념과 기능 중심 접근 방법으로 데이터 흐름 모델을 간단히 살펴보고 이 두 가지 모델을 조합하여 일관성 있고 완벽한 데이터베이스 설계 과정을 제시한다.

1.1 개체 관련성 모델

기존 구성 요소에는 독립적으로 존재하는 실세계의 개념적 대상으로서 개체, 개체들 사이에 존재하는 연관성(Association)을 나타내는데 속성, 그리고 개체들 간에 관련성을 나타내는 단항/이항/삼항 연관성이 있으며 개체 집합들 간의 대응성을 표현하기 위해 사상 원소수(1:1, 1:n, n:n)를 표현한다. 기본 구성 요소 외에도 개체 계층 구조에서 상위에 위치하는 개체는 일반화 개체가 되고 하위에 위치하는 개체는 세분화 개체가 되는 일반화 계층 구조와 단지 하나의 세분화 개체를 지니는 특별한 형태의 일반화 계층 구조로서 부분 집합, 단순 속성들의 집합으로 정의되는 복합 속성, 그리고 해당 개체와 관계를 가지는 개체가

없이는 존재할 수 없는 개체로서 약성 개체 등을 표현하여 데이터 중심의 모델링을 가능하게 한다. 그림 2는 개체 관련성 모델에서 사용되는 기호들이다.

〈그림 2〉 개체 관련성 모델에서 사용되는 기호

개념	개체	관련성	속성
표현			
개념	일반화 계층	부분집합	식별자
표현			

1.2 데이터 흐름 모델

구성 요소에는 정보를 생성, 사용, 조작, 및 제거함을 나타내는 시스템에 있어 행위자(Activity)를 표현되는 프로세스와 프로세스들 간에 정보의 교환을 표현하는 데이터 흐름, 정보의 저장소로서 데이터 저장, 그리고 시스템의 사용자를 표현하는 인터페이스 개념을 제공한다. 그림 3은 데이터 흐름 모델에서 사용되는 기호들이다.

〈그림 3〉 데이터 흐름 모델에서 사용되는 기호

개념	처리	흐름	데이터저장소	인터페이스
표현				

2. 개체 관련성 모델과 데이터 흐름 모델의 조합

개념적 설계와 기능적 분석을 함께 수행하므로 데이터베이스 설계의 일관성을 달성하기 위해 개체 관련성 모델에서 데이터에 대한 기술이 데이터 흐름 모델에서 프로세스에 의해 사용되어지거나 데이터 저장소에 유지되도록 데이터 흐름 모델을 확장한다. 확장된 모델에 의해 데이터와 기능 스키마를 생성하는 과정에서 상호 정교화와 완벽성 검사를 통해 상호 영향을 주게 된다. 개념적 데이터베이스 설계 단계에서 이루어지는 모든 작업들을 데이터 분석(Data Analysis)이라 하며 개념적 데이터베이스 스키마를 D-schema라 한다.

또한 기능 스키마를 F-schema라한다. 데이터

베이스를 위한 설계의 조합은 D-schema와 F-schema의 바로 전 스키마를 반영한 현 D-schema와 F-schema를 정교화 하는 것이다. 이러한 과정을 D-schema와 F-schema상에 반복함으로써 상호 일관성을 검사할 수 있다. D-schema와 F-schema의 조합에 의해 데이터와 기능을 분석하기 위한 접근 방법은 다음과 같은 과정에 의해 수행된다.

[분석방법]

- (1) 인터페이스를 식별한다.
- (2) 인터페이스와 시스템간에 입력과 출력 흐름을 식별한다.
- (3) 처음 골격 D-schema와 F-schema를 식별한다.
- (4)
 - (4.1) 바로 앞서 정의된 F-schema를 파악하면서, 앞서 정의된 D-schema로부터 D-schema를 정교화 한다.
 - (4.2) 바로 앞서 정의된 D-schema를 파악하면서, 앞서 정의된 F-schema로부터 F-schema를 정교화 한다.
 - (4.3) (선택 가능) D-schema와 F-schema의 상호 일관성 검사한다.
- (5) 요구 사항의 모든 개념이 스키마에 표현되지 않았다면 다시 (4)로 간다.
- (6) D-schema와 F-schema의 상호 완전성, 정확성, 판독성, 최소성, 함수적 독립성 등을 검사하여 요구되는 수준에 도달할 때까지 D-schema와 F-schema를 수정한다.

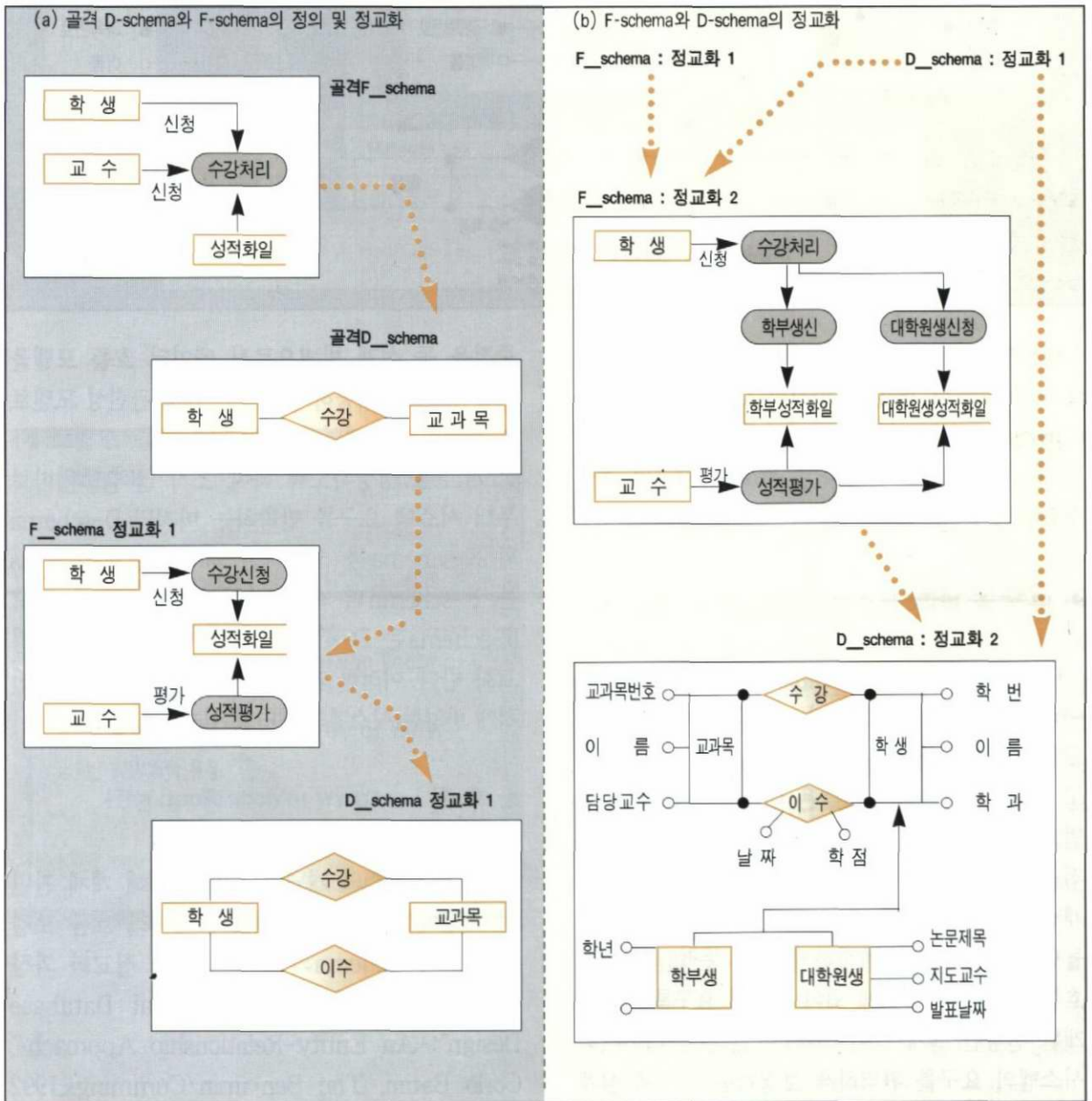
이러한 접근 방법은 설계자의 필요에 의해 변경될 수 있고, 제시된 접근 방법을 ‘학생의 학적 업무’와 관련하여 예로서 제시하면 그림 4와 같다.

이를 간단히 설명하면, 먼저 그림 4 (a)와 같이 골격 F-schema를 생성한다. 여기에는 수강 신청 업무의 프로세스만을 포함하며 수강 신청 업무를 수행하기 위한 학생과 교수 인터페이스를 포함한

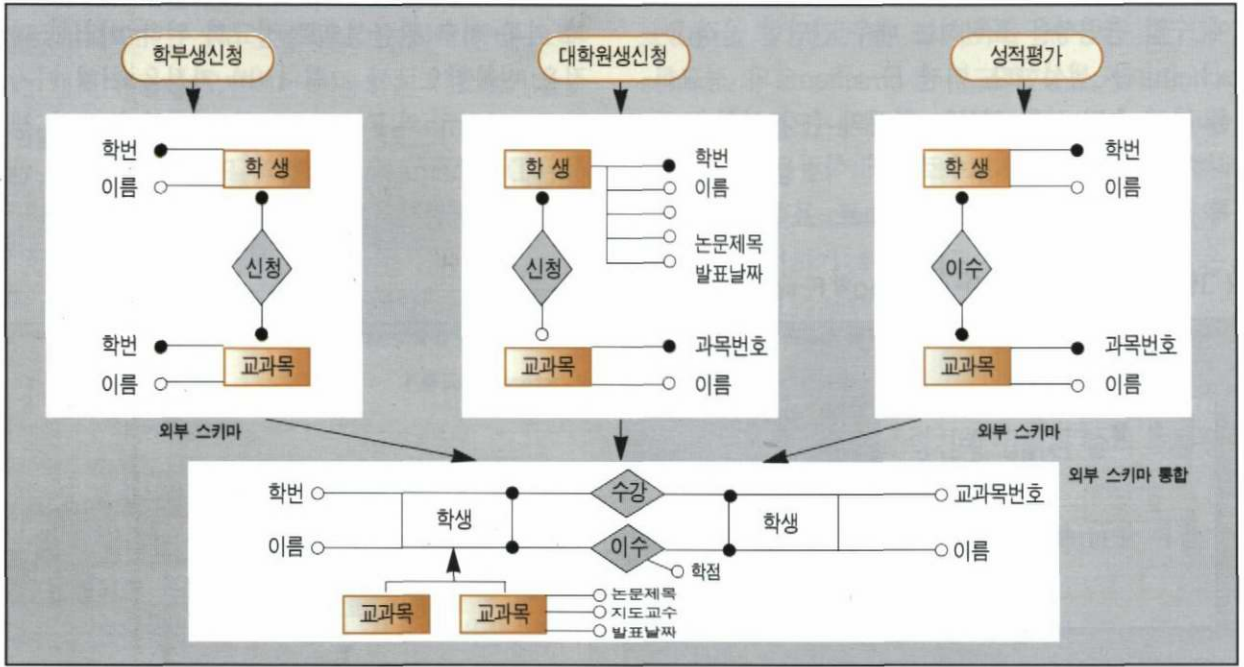
다. F-schema가 작성된 후, 학생과 교과목 사이에 수강 관련성을 표현하는 매우 간단한 골격 D-schema를 생성한다. 이젠 F-schema의 정교화를 통해 수강 신청 업무는 학생의 수강 신청을 관리하는 프로세스와 교수로부터 학점을 제공하도록 하는 성적 평가 프로세스로 표현된다. D-

schema는 다시 수강 관련성을 수강 신청 관련성과 과목 이수 관련성으로 정교화 된다. 이러한 과정을 반복함으로써 그림 4 (b) 과정을 거쳐 마지막 D-schema와 F-schema를 생성한다. 또한, 생성된 D-schema에 시스템의 요구가 완전하게 반영되었는지 검사하기 위한 과정은 그림 5와 같다.

< 그림 4 > 학사 업무에 관한 D-schema와 F-schema의 조합



〈 그림5 〉 완전성 검사



이러한 방법에 의해 골격 D-schema와 F-schema로부터 생성된 마지막 D-schema와 F-schema는 요구 사항을 완전하게 표현할 수 있으며 요구되는 수준의 판독성, 최소성, 함수의 독립성들을 얻을 수 있다.

3. 요약 및 결론

데이터베이스 설계를 위한 개념적 설계 도구로 여러 모델들이 제안되었다. 그 중에 개체 관련성 모델에 근거한 개념적 설계는 많은 영역에서 사용되고 있다. 개체 관련성 모델은 데이터 중심 접근 방법으로 데이터와 데이터의 특성에만 중점을 둔 접근 방법이다. 이러한 데이터 중심 접근 방법은, 개념적 설계를 위해 요구를 데이터 정적인 특성들만을 기본으로 모델링하게 되어 동적인 데이터 흐름을 표현할 수 없게 되어 시스템 요구를 완전하게 된 반영한 설계가 이루어지지 않는다. 데이터와 시스템의 요구를 완벽하게 보장하는 시스템 설계를 위해 개체 관련성 모델에 데이터보다 응용에

중점을 둔 설계 방법으로서 데이터 흐름 모델을 사용함으로써 가능하다. 이것은 개체 관련성 모델로 D-schema를 생성하고 데이터 흐름 모델로 F-schema를 생성하도록 하여, 초기 골격 스키마로부터 시스템 요구를 만족하는 마지막 D-schema와 F-schema를 생성하는 과정에서 D-schema는 F-schema의 의미를 반영하여 정교화 되고 F-schema는 D-schema의 의미를 반영하여 정교화 한다. 이러한 과정을 통해 시스템 요구를 완전하게 반영한 시스템의 개념적 설계가 가능하다.

4. 후기

데이터베이스의 개념적 설계를 위해 개체 관련성 모델에 의한 D-schema와 데이터 흐름 모델에 의한 F-schema의 조합에 의한 정교화 과정에 관한 많은 자료는 'Conceptual Database Design';-An Entity-Relationship Approach-, Corlo Batini, The Benjamin/Cummings,1992를 참조하였음을 밝힌다. 