
개념적 설계를 배제한 관계형 데이터베이스 스키마의 설계

엄 윤 섭*

A Design Methodology of Relational Database Schema Without the Conceptual Design Step

Yoon-Sup Um

본 연구는 2004학년도 경성대학교 학술지원 연구비에 의하여 연구되었음.

요 약

관계형 데이터베이스의 설계 과정은 요구사항 분석, ER 모델을 이용한 개념적 설계, 논리적 설계, 그리고 물리적 설계 과정으로 구성된다. 논리적 설계과정은 ER 다이어그램으로 구성된 개념적 스키마를 관계형 스키마로 전환하여 정규화시키는 과정으로 이루어진다. 이런 기존의 설계 과정에서 요구사항의 분석후 개념적 스키마를 도출하는 개념적 설계 과정이 명확치 않음으로 인해서, 실제 데이터베이스 설계현장에서 적용하기 어려운 점이 있다. 본 논문에서는 작업과정이 모호한 개념적 설계를 제거함으로 보다 체계적으로 데이터베이스 스키마를 설계할 수 있는 방법을 제안한다. 구조적 설계 방법을 이용하여 자료흐름도를 작성하고, 자료흐름도의 자료저장소에 파악된 애트리뷰트를 이용하여 관계합성으로 관계형 스키마를 도출한다. 관계형 스키마를 정규화하고, 정규화된 테이블로부터 데이터의 의미를 파악하여 테이블 관계도를 작성함으로써 최종 논리적 스키마를 결정한다.

ABSTRACT

The design process of a relational database system consists of requirement analysis, conceptual design using ER diagram, logical design, and physical design. In logical design process, the conceptual schema is transformed to relational schema, and relational schema is normalized. This traditional design process is hard to applied in real database design process, since there is an ambiguity in conceptual design process. In this paper, we suggest a new design process, which provides more structural design steps by removing the conceptual design process. In new approach, we produce the data flow diagram by the structural methodology. From the attributes in the data store of data flow diagram, we construct relational table schema, and we normalize relational schema. Finally we produced table relationship diagram in order to figure out relationships between tables.

1. 서 론

과거 40년 동안의 정보화 과정에서 데이터베이스 시스템은 가장 많이 사용되고 발전되어온 분야

중의 하나이다. 초기의 파일처리 방법에서 IMS로 대표되는 계층형 데이터모델과 DBTG로 표준화된 망형 모델을 거쳐서, Codd가 1970년대 초에 제안한 관계형 데이터 모델[Codd]은 1970년대와 1980

년대에 부단히 발전하여 많은 데이터베이스 관련 기술을 개발하는데 큰 역할을 하였다. 1990년대에 들어 의미적 데이터 모델(semantic data model)에서 시작한 데이터 모델링 기법의 발달로 객체지향 데이터 모델이 개발되어 상용화 되고 있다. 그러나 지금까지 대부분의 데이터베이스 시스템의 구축은 관계형 데이터베이스로 이루어지고 있기 때문에, 현재는 관계형 데이터베이스 시스템의 시대라고 볼 수 있다. 관계형 데이터 모델의 데이터베이스의 설계에서는 자료수집과 요구사항 분석 후 일반적으로 ER 모델[Chen]을 이용하여 개념적 설계를 한 후, 논리적 설계단계에서 관계형 데이터베이스의 스키마가 완성된다. 이런 설계과정은 대부분의 데이터베이스 저서에서 기술되어 있지만, 실제 데이터베이스 설계에서 적용하기 어렵기 때문에 잘 사용되지 않고 있다[batini]. 실제의 설계에서는 알려진 기존의 설계방법보다는 주로 설계자의 개인적 경험에 의하여 데이터베이스가 설계되고 있으므로, 현장에 적합한 설계방법이 개발이 필요하다.

본 논문에서는 기존의 설계방법에 대한 문제를 논하고, 문제점을 제거하기 위한 새로운 관계형 데이터베이스 설계방법을 제시한다. 제안되는 설계방법에서는 구조적 방법[3,4,13]을 이용하여 시스템 요구사항의 분석과 설계를 수행한다. 구조적 기법의 결과로 작성된 자료흐름도를 이용하여 데이터베이스 스키마를 설계한다. 자료흐름도 내의 자료저장소로부터 데이터베이스의 스키마를 설계하고, 처리기를 이용하여 응용 프로그램을, 그리고 자료흐름선을 이용하여 데이터베이스 시스템에서 필요한 입출력을 설계한다. 스키마의 설계는 자료저장소의 애트리뷰트를 이용하여 초기 테이블을 도출하고, 함수적 종속성을 이용하여 키를 찾아내고 정규화를 수행한다. 정규화된 테이블로부터, 각 테이블 사이의 관계를 파악하여 테이블 관계도를 작성한다. 테이블 관계도는 기존의 시스템 설계에서 ER 스키마와 같으며, 데이터베이스의 무결성을 설정하는 기초 자료로 사용된다.

본 논문에서 제안하는 데이터베이스의 설계방법의 특징은 기존 설계방법에서 체계적으로 수행하기 힘들었던 개념적 설계를 배제하고, 지금까지 잘 발달된 여러 가지 설계 기법을 재구성하여 보다 체계적이고 합리적인 단계를 거쳐서 데이터베이스를 설계하는 방법을 제시하는 데 있다. 또한 데이터베이스의 설계를 데이터 위주 설계 방법이나 기능위주 설계 방법으로 구분하여 선택적인 방법으로 수행하는 것을 탈피해서, 두 가지 방법을 통합적으로 동시에 수행하는 것이다. 실제로 정보 시스템을 어느 한쪽 방향에 치우쳐 설계하는 것은 장단점을 모

두 가짐으로 바람직하지 못하기 때문에, 본 논문에서는 두 가지 방법을 동시에 통합할 수 있는 방법을 제안한다.

2. ER 모델을 이용한 개념적 설계의 문제점과 관련 연구

데이터베이스를 설계할 때 요구사항 분석 후 ER 모델을 이용하여 개념적 스키마를 설계하는 것이 일반적이다. 그러나 실제 현장에서는 ER 모델을 이용하는 개념적 스키마를 설계하는 경우는 별로 없다고 알려져 있다[7]. 그 이유는 첫째, ER 다이어그램을 작성하는 사고과정이 자연스럽지 못한 점이다. 시스템 분석 단계에서 수집한 자료로써 ER 다이어그램을 그리는 것보다 관계형 스키마를 나타내는 형식인 테이블로 그리는 것이 자연스럽다. 예를 들어 교수와 학과의 관계를 살펴보자. 표 1에서 교수 테이블과 학과 테이블에서 교수가 속한 학과를 하나의 애트리뷰트로 나타내는 것은 자연스럽다. 그러나 그림 1의 ER 다이어그램을 그리기 위해서는 교수 엔티티와 학과 엔티티 사이에는 소속학과라는 또 하나의 독립된 개념을 생각해야 하는 부담이 생긴다. 더구나 두 엔티티 사이의 소속학과를 관계성(relationship)의 개념으로 생각하는 것은 보다 추상적이어서 익숙하지 않을 경우에는 사고의 부담이 크다.

그림 1. 교수와 학과를 나타내는 테이블

Table 1. Table for Professor and Department

(a) 교수 테이블

교수번호	교수이름	전공	직급	학과
8923	이휘소	핵이론	정교수	S01
7658	김진기	신학	전임강사	G02
9453	박철수	반도체	부교수	E12
2617	정두한	모바일	조교수	E17

(b) 학과 테이블

학과번호	학과이름	정원
S01	물리학과	20
G02	신학과	25
E12	전자과	45
E17	통신과	55

둘째, ER 다이어그램은 보통 데이터를 기술한 문장으로부터 작성되는 점이다. 따라서 ER 다이어

그램을 그리기 위해서는 사용기관에서 필요한 데이터를 문장으로 기술해야 한다. 여기서 문제는 필요한 데이터를 문장으로 기술하는 데는 너무 많은 시간과 노력이 필요할 뿐만 아니라, 작성된 문장의 모호성도 문제가 될 수 있다.

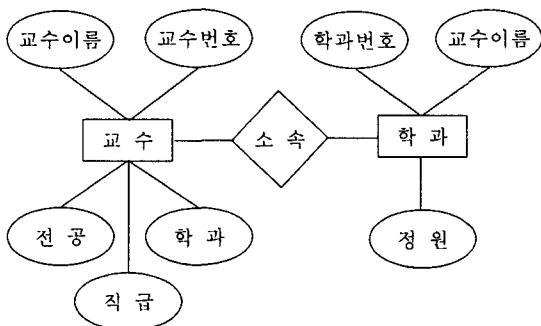


그림 1. 교수와 학과 관계의 ER 다이어그램

Fig. 1. ER Diagram for Professor and Department

셋째, 구조적 방법을 이용하여 요구사항을 분석 또는 설계할 경우, 자료흐름도로부터 ER 다이어그램을 작성할 수 있는 체계적인 알고리즘이 존재하지 않는다. 구조적 방법으로 업무분석을 한 후 ER 모델로 개념적 설계를 한 경우는[7, 13], 자료흐름도 작성 이후 논리적인 단계를 거치지 않고 별도로 작성된 ER 스키마가 설명된다.

이러한 여러 가지 이유 외에도 ER 모델의 근본적인 문제점은 ER 모델에서 사용하는 관계성(relationship)은 Codd가 관계형 데이터 모델을 제안했을 때 이미 관계(relationship)는 ER 모델에서의 엔티티와 관계성(relationship)을 모두 나타낼 수 있다[10]. 따라서 데이터베이스 설계시 ER 모델을 사용하지 않고 관계형 테이블을 사용하더라도 스키마의 설계에는 문제가 발생하지 않는다.

데이터베이스 설계 관련 저서에서 설계 절차를 살펴보면, 거의 대부분 ER 모델을 기반으로 데이터베이스를 설계했다[1,2,5,6,7]. Kendall[13]은 구조적 기법으로 업무분석을 하고 자료흐름도를 작성한 후, ER 다이어그램을 그려서 개념적 설계를 하는 것으로 기술되어 있다. 그러나 자료흐름도로부터 논리적으로 ER 스키마를 도출하는 방법이 기술되어 있지 않다. Batini[7]는 자료흐름도로 기능분석을 한 다음 바로 ER 다이어그램을 설명하고, 둘 사이의 보완적인 기능을 이용하여 자료흐름도와 ER 다이어그램을 정제하지만, 기능 분석 후 합리적

인 과정없이 갑자기 ER 다이어그램을 도출하는 심각한 논리적 비약이 있다. 문송천[2]은 구조적 기법에서 사용하는 자료흐름도는 DB 접근방식에서는 더 이상 적합하지 않다고 결론짓고, ER 모델을 중심으로 데이터베이스 스키마를 설계했다. 이준식[5]은 프로세스 모델링과 ER 모델링을 상호보완적으로 병행하여 데이터 스키마의 설계를 시도했다. 조규익[6]은 기능중심적 모델링과 데이터 중심적 모델링 가운데 데이터 중심적 설계의 융통성을 장점으로 ER 모델을 기반으로 설계했다. Hernandez[12]는 현장에서 쌓은 경험을 중심으로 체계적인 정규화 과정이나 개념적 설계 없이 관계테이블을 설계했으나, 저서에서는 ER 모델을 따랐다고 기술했다.

3. 개념적 설계를 생략한 논리적 스키마 설계 방법

본 논문에서는 ER 모델을 이용한 개념적 설계의 문제점을 해결하기 위해서, 개념적 설계 과정을 배제하고 데이터베이스 설계를 할 수 있는 방법을 제시한다.

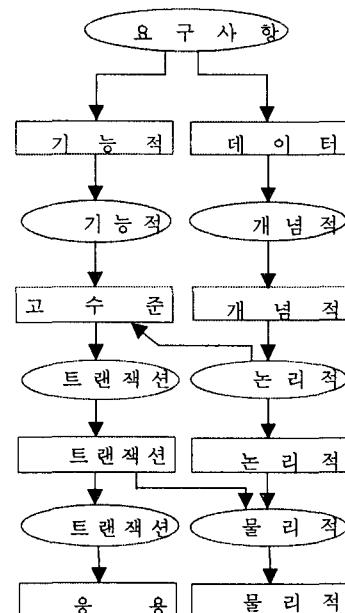


그림 2. 기존의 데이터베이스 스키마 설계절차
Fig. 2. Traditional Design Process for Database Schema

그림 2는 보편적으로 알려진 데이터베이스 스키마의 설계 절차이다. 여기서는 개념적 설계, 논리적 설계, 그리고 물리적 설계의 순서 스키마의 설계가 진행된다. 그림 3은 본 논문에서 제안하는 새로운 데이터베이스 스키마 설계방법이다. 먼저 시스템 분석과 설계 방법으로는 일반적으로 많이 사용되는 구조적 기법으로 작성된 자료흐름도의 자료저장소에는 시스템이 필요로 하는 모든 애트리뷰트가 도출된다. 이 애트리뷰트를 관계합성(relational synthesis)을 이용하여, 관계형 데이터베이스 스키마를 도출한다. 관계합성이란 필요한 애트리뷰트를 파악한 후, 애트리뷰트간의 종속성을 고려하여 관계형 데이터베이스의 스키마를 설계하는 방법이다[11]. 따라서 기존의 ER 모델을 이용한 개념적 설계를 수행하지 않는 것이다. 대신 스키마가 도출된 후 ER 스키마와 유사한 테이블 관계도를 이용하여 데이터 사이의 관계를 파악한다. 이 설계 과정에서는 자료흐름선을 이용하여 입출력 설계를 할 수 있으며, 처리기를 기반으로 트랜잭션을 설계할 수 있다. 본 논문에서는 논리적 스키마의 설계 과정의 기술에 중점을 둔다.

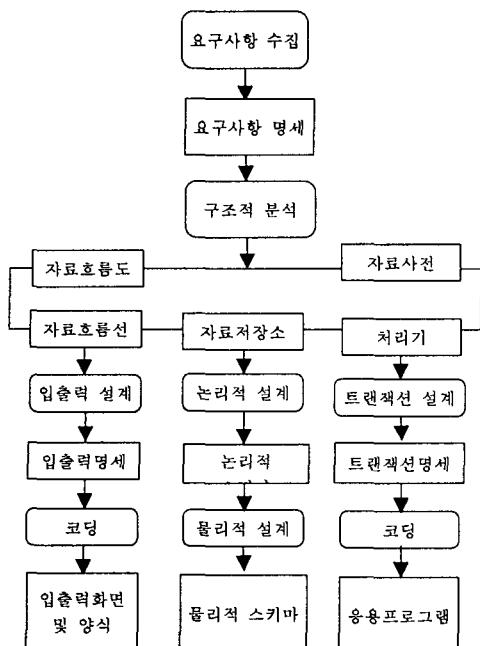


그림 3. 관계합성에 의한 데이터베이스 스키마의 설계 절차

Fig. 3. Database Design Process Using Relational Synthesis

4. 구조적 기법을 이용한 시스템 분석과 설계

일반적으로 데이터베이스를 설계하는 방법으로는 데이터 중심 접근(data-driven approach)방법과 기능 중심 접근방법(function-driven approach)

이 많이 알려져 있다. 데이터 중심 접근 방법은 설계에서 데이터와 데이터의 속성에 초점을 맞추어, 먼저 데이터베이스를 설계한 다음 데이터베이스의 용용 프로그램을 설계한다. 이 방법에서는 보통 ER 모델을 이용하여 데이터베이스를 설계한다.

기능중심 설계방법은 데이터보다는 시스템의 기능에 초점을 맞춘 설계방법으로, 일반적으로 구조적 기법을 이용한다. 구조적 기법은 1969년 Dijkstra가 순차, 선택, 반복의 세 가지 기본 제어로 goto 문의 사용 없이 모든 프로그램의 작성이 가능한 것을 증명함으로 도입된 구조적 프로그래밍의 원리를 소프트웨어 개발 방법으로 도입함으로 사용되기 시작했다. 구조적 설계기법은 업무 수행에서 사용되는 자료의 흐름과 흐름에 따른 가공 절차에 따라서 시스템을 설계하는 방법이다. 구조적 기법을 이용한 설계에서는 자료흐름도, 자료사전, 미니명세서가 작성된다. 자료흐름도는 외부관련자, 자료흐름선, 처리기, 그리고 자료저장소로 구성된다.

데이터베이스의 설계에서 기능의 파악없이 데이터만의 도출이 사실상 쉽지 않으며, 업무의 파악에서 기능과 데이터가 동시에 도출된다는 점에서 본 연구에서는 구조적 기법을 이용한 기능중심 접근방법을 사용한다. 구조적 기법으로 작성되는 자료흐름도의 자료저장소에는 데이터베이스 시스템에 필요한 모든 애트리뷰트가 파악되어 있기 때문에, 데이터중심 접근방법에서와 마찬가지로 데이터에 초점을 맞출 수 있다. 따라서 구조적 기법을 이용하여 데이터와 기능 양쪽 모두를 고려하여 시스템을 설계할 수 있다.

본 논문에서 새로운 데이터베이스의 설계 방법을 설명하기 위한 예로 커피가공회사의 원료 구매와 제품 판매를 관리하기 위한 데이터베이스 시스템을 설계한다. 시스템의 배경도는 그림 4와 같다. 외부 인터페이스로는 공급자, 고객, 생산부, 운송자, 회계과가 있다. 공급자는 커피원두를 공급하고, 고객은 가공된 커피믹스 제품을 구입한다. 운송자는 커피원두나 커피믹스 제품을 운반하는 관련자이며, 회계과에서는 구입한 커피원두의 가격을 지불한다.

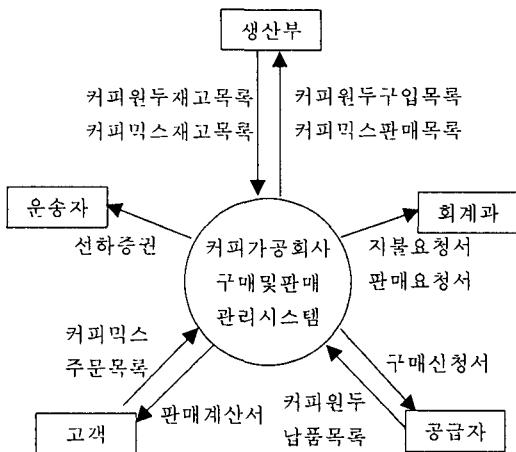


그림 4. 배경도

Fig. 4. Context Data Flow Diagram

생산부는 커피원두로 커피믹스 제품을 생산하는 역할을 합니다. 본 시스템에서는 생산부의 세부 기능은 생략하고, 단지 원두의 소모나 제품의 재고만을 보여주는 역할을 합니다. 커피원두의 재고는 매주 한번씩 체크되어, 재고량이 일정량 이하로 떨어지면 원두구매가 발주된다. 구매와 판매관리 시스템의 주요업무는 구매업무와 판매업무로 그림 5와 같이 개략도로 나타낼 수 있다.

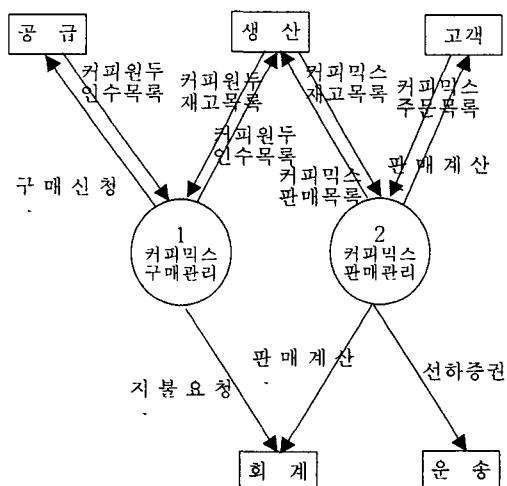


그림 5. 개략도

Fig. 5. Overview Data Flow Diagram

설명을 간략화하기 위해서, 구매업무에 관한 부분만으로 데이터베이스 스키마를 설계한다. 구매업무의 최하위도는 그림 6과 같이 작성할 수 있다. 생산부의 커피원두 재고량을 살펴보아, 구매발주가 필요할 경우 커피원두목록 파일과 공급자 파일로부터 필요한 데이터를 검색하여 구매신청서를 작성한다. 작성된 구매신청서는 구매신청 파일에 저장됨과 동시에 공급자에게 보내진다. 공급자가 커피원두를 납품하면, 납품목록을 구매신청 파일과 비교하여 문제가 없을 경우 인수보고서를 작성한다. 작성된 인수보고서는 인수내역 파일에 저장하는 동시에, 생산부의 커피원두 재고목록을 갱신한다. 인수내역 파일로부터 지불요청서를 작성하여, 지불요청내역 파일에 저장하는 동시에 회계파일에 보내어 대금지불을 요청한다. 공급자 파일은 새로운 공급자가 나타날 경우나 기존 공급자가 없어질 경우 공급자 관리 처리기에서 파일이 수정된다. 커피원두목록 파일의 수정도 동일한 방법으로 수행된다.

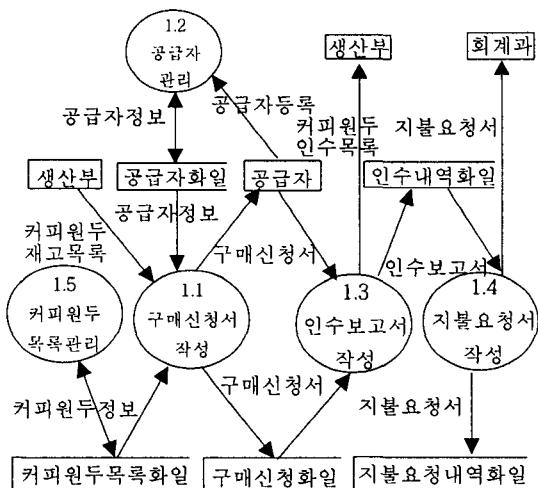


그림 6. 구매업무 최하위도

Fig. 6. Lowest-Level DFD for Purchasing

자료 흐름도는 계층적으로 작성된다. 상위도는 개념적으로 작성되고, 중간계층도에서 상세화해 나가면서, 최하위도는 더 이상 분할할 수 없는 단위기능의 처리기로 구성된다. 그림 4의 배경도는 시스템의 목적과 범위가 명시되며, 관련자를 이용하여 외부와의 경계를 표시한다. 개략도에서는 시스템의 주요업무를 분할한다. 그림 5의 개략도에서 구매업무와 판매업무가 주요 업무로 명시되었다. 그림 6은 구매업무의 최하위도이다. 자료흐름도에

서 처리기로부터 출력되는 자료흐름선은, 출력자료에 대응되는 입력 자료가 처리기로 들어와야 한다. 이것을 자료보존의 법칙이라고 한다. 본 설계 방법에서는 자료저장소를 이용하여 관계합성을 통한 데이터베이스를 설계하기 때문에 자료보존의 법칙을 다음과 같이 수정한다.

- 모든 입력 자료는 자료저장소에 저장되어야 하고, 모든 출력자료는 자료저장소로부터 산출할 수 있어야 한다.

자료흐름도에 따른 자료사전은 다음과 같다.

커피원두재고목록 = 일자 + 커피원두 + 재고량
커피원두목록 = {커피원두}

커피원두=원두코드+원두이름+원두종류

+원두단가

원두종류 = [레귤러 | 디카페인]

공급자화일 = {공급자}

공급자 = 공급자코드+공급자상호+공급담당자
+공급자주소 +공급자우편번호

공급자코드 = *회사 내부에서 자체적으로 설정
된 공급자 식별코드*

구매신청 화일 = {구매신청서}

구매신청서=구매신청번호+공급자+구매신청일
자 +수령예정일자+[커피원두+주문량
+소계] +구매총액

커피원두인수목록 = *공급자로부터 접수한
선적내역 *

={커피원두+인수량+소계}

인수내역화일 = {인수보고서}

인수보고서 = 인수번호+구매신청번호+선적번호
+인수담당자+인수일자+운송자
+공급자 + {커피원두+인수량
+상자수+소계}+인수총액

지불요청내역화일 = {지불요청서}

지불요청서 = *회계과로 보내는 구매원두의
지불요청 양식*
=지불요청번호+구매신청번호+
(선적번호+인수번호)+공급자
+지불요청일자+ {커피원두+인수량
+소계} +총액+운송비+총비용

5. 자료 저장소를 이용한 관계형 데이터베이스 스키마의 설계

관계형 데이터베이스의 대표적인 두 가지 설계

방법은 하향식 설계(top-down design)와 관계합성 (relational synthesis)이다. 하향식 설계는 ER 모델과 같은 고수준의 개념적 스키마를 설계한 후, 관계형 스키마로 변형하는 방법이다. 관계합성은 데이터베이스의 애트리뷰트를 먼저 추출하고, 애트리뷰트 사이의 함수적 종속성을 파악하여 관계형 데이터베이스를 설계하는 방법이다. 하향식 설계방법의 문제는 ER 모델이 실제의 데이터베이스 설계에서 별로 사용되지 않고 있으며, 실제의 설계에서 많은 경우 개념적 스키마의 설계 없이 바로 논리적 설계에 들어간다. 따라서 본 논문에서는 하향식 설계방법을 사용하지 않고, 자료저장소를 구축하면서 파악된 애트리뷰트를 이용하여 관계합성 방법으로 데이터베이스 스키마를 설계한다.

5.1 테이블의 구성

데이터베이스 테이블은 데이터 테이블, 링크 테이블로 구분할 수 있다. 데이터 테이블은 ER 모델에서 엔티티 테이블에 해당하고, 링크 테이블은 두 엔티티 사이의 관계를 맺어주는 관계성 테이블에 해당한다. 먼저 자료저장소에서 엔티티가 될 수 있는 애트리뷰트를 모아 데이터 테이블을 작성하고, 사용기관의 업무내용을 파악하여 링크 테이블을 작성한다. 위의 구매업무를 위한 자료흐름도에는 다섯 개의 자료저장소가 있다. 자료저장소의 애트리뷰트로 테이블을 구성하면 다음과 같다

테이블	애트리뷰트
커피원두	원두코드, 원두이름, 원두종류, 원두단가
공급자	공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호
구매신청	구매신청번호, 구매신청일자, 수령예정일자, 공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호, {원두코드, 원두이름, 원두종류, 원두단가, 주문량, 소계}, 구매총액
인수내역	인수번호, 구매신청번호, 선적번호, 인수담당자, 인수일자, 운송자, 공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호, {원두코드, 원두이름, 원두종류, 원두단가, 인수량, 소계}, 총액, 운송비, 인수총액
지불내역	지불요청번호, 지불요청일자, {공급

	자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호, 구매신청번호, 인수번호, 선적번호, 원두코드, 원두이름, 원두종류, 원두단가, 인수량, 소계), 총액, 운송비, 지불총액)
--	---

5.2 정규화

각 테이블의 모든 애트리뷰트가 원자값을 가질 때 테이블은 제 1정규형이 된다. 구매신청 테이블, 인수내역 테이블, 지불내역 테이블에는 집합으로 명시된 반복그룹이 존재한다. 제 1정규화를 위해서 반복그룹을 제거해야 한다. 또한 구매신청 테이블의 소계와 구매총액, 인수내역 테이블의 소계, 총액, 운송비, 그리고 인수총액, 지불내역 테이블의 소계, 총액, 운송비, 지불총액은 추출 애트리뷰트들이다. 반복그룹과 추출 애트리뷰트를 제거하여 다음과 같은 제 1정규형의 테이블을 만든다.

테이블	애트리뷰트
커피원두	원두코드, 원두이름, 원두종류, 원두단가
공급자	공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호
구매신청	구매신청번호, 구매신청일자, 수령예정일자, 공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호
구매목록	구매신청번호, 원두코드, 원두이름, 원두종류, 원두단가, 주문량
인수내역	인수번호, 구매신청번호, 인수담당자, 인수일자, 운송자, 선적번호, 공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호
인수목록	인수번호, 원두코드, 원두이름, 원두종류, 원두단가, 인수량
지불내역	지불요청번호, 지불요청일자
지불목록	지불요청번호, 공급자코드, 공급자상호, 공급담당자, 공급자주소, 공급자우편번호, 선적번호, 인수번호, 구매신청번호, 원두코드, 원두이름, 원두종류, 원두단가, 인수량

커피원두 테이블과 공급자 테이블에서의 함수적 종속성은 다음과 같다.

원두코드 -> 원두이름, 원두종류, 원두단가
공급자코드 -> 공급자상호, 공급담당자, 공급자주소, 공급자우편번호

따라서 이 두 테이블의 기본키는 각각 원두코드와 공급자코드가 된다. 구매신청 테이블에서 추이적 종속성을 제거하기 위해서 공급자 관련사항을 제거하면, 함수적 종속성은 다음과 같다.

구매신청번호 -> 공급자코드, 구매신청일자, 수령예정일자,

따라서 다음과 같은 구매신청테이블이 도출된다.

구매신청 (구매신청번호, 구매신청일자, 수령예정일자, 공급자코드)

구매목록 테이블의 함수적종속성은 다음과 같다.

구매신청번호, 원두코드 -> 주문량

원두코드 -> 원두이름, 원두종류, 원두단가

여기는 부분종속성이 존재하기 때문에 원두관련 애트리뷰트를 제거해야 한다. 원두관련 애트리뷰트는 이미 커피원두 테이블에 분리되어 있다. 인수내역 테이블의 공급자 관련 애트리뷰트는 구매신청 테이블의 경우와 마찬가지로 분리되고, 나머지 애트리뷰트의 함수적 종속성은 다음과 같다. 구매신청번호가 공급자코드를 결정하기 때문에 공급자 정보는 제거된다.

인수번호 -> 구매신청번호, 인수담당자,

인수일자, 운송자, 선적번호

인수내역 테이블의 함수적 종속성은 구매내역과 유사하며, 다음과 같다.

인수번호, 원두코드 -> 인수량

지불내역 테이블의 정규화는 지불요청번호가 지불요청일자를 결정하기 때문에 다음 함수적 종속성이 성립한다.

지불요청번호 -> 지불요청일자

지불목록 테이블의 정규화에서는 선적번호와 구매신청번호, 공급자코드, 원두코드는 인수번호에 의해서 결정되고, 남은 애트리뷰트는 지불요청번호와 인수번호뿐이다. 결과적으로 다음과 같은 제 3정규형의 테이블을 산출할 수 있다.

5.3 테이블 관계도 작성

기존 ER 모델을 이용한 개념적 스키마는 관계 데이터베이스 스키마를 설계하기 이전 단계로 사용되었다. 본 논문에서는 이 과정을 생략하고 자료 흐름도의 자료저장소로부터 관계 스키마를 관계 합성에 의해서 설계했다. 스키마를 생성한 다음 단계는, 각 테이블 사이의 관계를 파악하여, 테이블 사이의 외래키를 찾아서 참조무결성 제약조건을 설정하고, 테이블 사이의 참여차수(degree of participation)와 참여비율(cardinality ratio)를 알아보아야 한다. 이를 위해서 이미 도출된 8개의 테이블을 하나씩 차례로 비교하여 각각의 두 테이블 사이의 관계를 알아낸다. 결과는 그림 7과 같다.

공급자 테이블의 키인 공급자코드는 구매신청 테이블에서 외래키가 되며, 1:N의 관계를 가진다. 공급자 테이블의 참여차수(1,1)은 구매신청 레코드는 최소, 최대 하나의 공급자 레코드에 연관되어 있다는 의미이고, 구매신청 테이블의 참여차수 (0, N)은 공급자 테이블의 각 레코드는 구매신청 레코드와 연관이 없을 수도 있고, 다수의 레코드와 연관이 있을 수 있음을 나타낸다. 공급자 테이블의 R은 restrict 제약을 나타낸다.

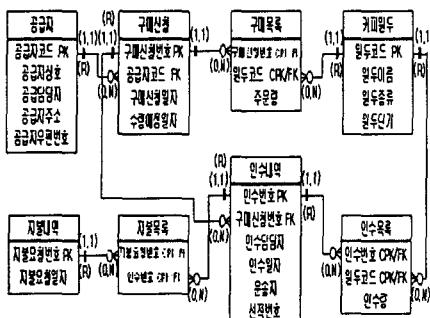


그림 7. 원두 구매 업무를 위한 테이블 관계도
Fig. 7. Table-Relationship Diagram for Coffee Bean Purchasing

테이블 관계도는 ER 다이어그램을 그리는 것과 유사하다. 테이블 관계도의 작성이 필요한 것은 물리적 설계에서 참조무결성이나 데이터 삽입, 삭제, 생성시 부과되는 데이터에 관한 제약을 설정하기 위해서이다. 이 방법이 기존의 방법에 대한 이점은 데이터베이스의 설계 작업이 자료흐름도로부터 테이블 관계도를 작성하는 과정이 보다 기계적인 단계에 의해서 진행된다는 점이다. 여기서는 요구사항 분석후 갑자기 ER 다이어그램을 그리는 비약없이, 자료저장소로부터 테이블 관계도까지 논리적으로 작업을 진행할 수 있다.

6. 응용 프로그램과 인출력 설계

스키마의 설계가 끝나면 처리기로부터 트랜잭션을 설계할 수 있다. 트랜잭션의 설계는 데이터베이스 액션 다이어그램(database action diagram)으로 명시한다. 구매업무와 관련해서 다음과 같은 5개의 처리기가 있고, 각 처리기는 하나의 응용 프로그램에 해당한다.

- 가. 공급자관리
 - 나. 구매신청서 작성
 - 다. 인수보고서 작성
 - 라. 지불요청서 작성
 - 마. 커피원두 목록관리

자료흐름도에서 모든 자료의 흐름은 자료흐름선으로 표시되므로, 자료흐름선으로부터 입출력화면을 설계할 수 있다. 자료흐름선이 처리기로 들어갈 경우 입력선이 되고, 처리기로부터 나을 경우 출력선이 된다. 처리기로의 자료 입력은 외부의 관련자나 내부의 자료저장소로부터 입력된다. 출력은 화면 또는 보고서 형식을 통해서 이루어진다. 구매신청서 작성을 위한 입력 데이터는 공급자 테이블과 커피원두 테이블로부터 도출할 수 있다. 공급자 테이블의 대부분의 애트리뷰트와 재고량이 일정량이 하인 커피원두에 관한 정보, 그리고 각 원두의 주문량, 구매신청일자, 수령예정일자로 구성될 것이다. 공급자에게 보내는 구매신청서에는 각 커피원두의 주문 내역으로 구성된다. 공급자에 관한 정보는 공급자 테이블로부터 가져오고, 커피원두에 관한 정보는 커피원두 테이블에서 추출할 수 있다. 구매신청번호, 구매일자, 수령일자, 원두단가, 주문량은 작성자가 화면을 통해서 입력한다.

7. 결론

본 논문에서는 기존의 데이터베이스 설계 방법인 ER 모델을 이용한 개념적 설계를 거친 다음 논리적 설계에서 데이터베이스 스키마를 도출하는 방법의 문제점을 지적하고, 새로운 데이터베이스 스키마 설계방법을 제시했다. 기존의 개념적 설계에서는 ER 스키마를 설계하기 위해서 필요한 데이터를 문장으로 기술하는 작업의 어려움과 ER 디자인 그램 작성 시 엔티티와 관계성을 도출할 때 논리적인 방법이 없기 때문에 실제 현장에서 사용하기 어려운 문제점이 있다. 본 논문에서 제안하는 설계 방법에서는 기존에 잘 개발된 구조적 설계방법을 이용하여 단계적으로 관계형 데이터베이스 스키마를 설계할 수 있다. 먼저 구조적 설계방법을 적용하여 자료흐름도, 자료사전, 미니명세서를 작성한다. 자료저장소로부터 관계합성을 위해서 테이블을 도출하여 정규화하고, 각 테이블 사이의 관계를 파악하여 테이블 관계도를 작성한다. 자료흐름선으로부터 데이터베이스 시스템의 입출력을 설계하고, 미니명세서로부터 응용 프로그램을 설계할 수 있다. 새로운 방법을 사용함으로써, 시스템의 설계에서 가능한적인 설계와 데이터 중심적인 설계를 동시에 수행할 수 있을 뿐만 아니라, 소프트웨어 공학에서 개발된 구조적 설계방법을 데이터베이스 설계에 도입함으로써, 논리적인 비약 없이 단계적으로 스키마를 도출할 수 있는 이점이 있다.

참고문헌

- [1] 나민영, 데이터베이스 설계, 기한재, 1996.
- [2] 문송천, Dataware 설계론, 아이포스트, 1999.
- [3] 박재년, 구조적 시스템 분석과 설계, 정의사, 1992.
- [4] 우치수, Software 공학, 상조사, 1994.
- [5] 이춘식, 데이터베이스의 설계와 구축, 한빛미디어, 2002.

- [6] 조규익, 데이터베이스 설계, 홍릉과학출판사, 1994.
- [7] Batini, Conceptual Database Design, Benjamin/Cummings, 1992.
- [8] Chen, P. [1976] "The Entity Relationship Model-Toward a Unified View of Data," TODS, 1/1, pp. 9-36, March 1976.
- [9] Codd, E.[1970]"A Relational Model for Large Shared Data Banks," CACM, 13/6, pp. 377-387, June 1970.
- [10] Date, C. "Entity/Relationship Modeling and the Relational Model." in C. Date and Hugh Darwen, Relational Database Writings 1989-1991, Mass.; Addison-Wesley, 1992.
- [11] Elmasri, R. & Navathe, S. Fundamentals of Database Systems, Addison-Wesley, 2000.
- [12] Hernandez, M. Database Design for Mere Mortals, Addison-Wesley, 1997.
- [13] Kendall, P. Introduction to Systems Analysis & Design: A Structured Approach IRWIN, 1994.

저자소개



엄윤섭 (Yoon-Sup Um)

1979년 부산대학교 졸업
 1987년 아리조나 주립대학 컴퓨터공학과 (공학석사)
 1990년 아리조나 주립대학 컴퓨터공학과 (공학박사)
 1991년 ~ 현재. 경성대학교 컴퓨터공학과 교수
 연구분야: 데이터베이스, 객체지향 시스템