

자료흐름도에 기반한 관계형 데이터베이스의 설계방법

엄운섭 권중장 정희석 김창동
경성대학교 컴퓨터공학과

A Design Methodology of Relational Database Based on Dataflow Diagram

Yoonsup Um, Jung Jang Kwon, Hee-Suk Jung, Chang-Dong Kim
Dept. of Computer Engineering, Kyung Sung University
E-mail : ysum@ks.ac.kr

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.

키워드

데이터베이스 설계, 관계형 데이터베이스, 구조적 설계, ER 모델, 정규화

1. 서 론

관계형 데이터 모델의 데이터베이스의 설계에서는 자료수집과 요구사항 분석 후 일반적으로 ER 모델[4]을 이용하여 개념적 설계를 한 후, 논리적 설계단계에서 관계형 데이터베이스의 스키마가 완성된다. 이런 설계과정은 대부분의 데이터베이스 저서에 기술되어 있지만, 실제 데이터베이스 설계에서 적용하기 어렵기 때문에 잘 사용되지 않고 있

다[3]. 실제의 설계에서는 알려진 기존의 설계방법 보다는 주로 설계자의 개인적 경험에 의하여 데이터베이스가 설계되고 있으므로, 현장에 적합한 설계방법이 개발이 필요하다.

본 논문에서는 기존의 설계방법에 대한 문제를 살펴보고, 문제점을 개선하기 위한 새로운 관계형 데이터베이스 설계방법을 제시한다. 제안되는 설계방법에서는 구조적 방법[1,2]을 이용하여 시스템 요구사항의 분석과 설계를 수행한다. 구조적 기

법의 결과로 작성된 자료흐름도의 자료저장소로부터 데이터베이스의 스키마를 설계한다. 스키마의 설계는 자료저장소의 애트리뷰트를 이용하여 초기 테이블을 도출하고, 함수적 종속성을 이용하여 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 모두 가짐으로 바람직하지 못하기 때문에, 본 논문에서는 두 가지 방법을 동시에 통합할 수 있는 방법을 제안한다.

II. ER 모델을 이용한 개념적 설계의 문제점

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[3]. 그 이유로 생각할 수 있는 것은 ER 다이어그램을 작성하는 사고과정이 자연스럽지 못한 점이다. 시스템 분석 단계에서 수집한 자료로써 ER 다이어그램을 그리는 것보다 관계형 테이블로 그리는 것이 자연스럽다. 예를 들어 교수와 학과의 관계를 살펴보자. 그림 1에서 교수 테이블과 학과 테이블에서 교수가 속한 학과를 하나의 애트리뷰트로 나타내는 것은 자연스럽다. 그러나 그림 2의 ER 다이어그램을 그리기 위해서는 교수 엔티티와 학과 엔티티 사이에는 소속학과라는 또 하나의 독립된 개념을 생각해야 하는 부담이 생긴다. 더구나 두 엔티티 사이의 소속학과를 관계성(re-relationship)의 개념으로 생각하는 것은 추상적이어서 익숙하지 않을 경우에는 사고의 부담이 크다.

다른 이유로는 ER 다이어그램은 보통 데이터를 기술한 문장으로부터 작성되는 점이다. 따라서 ER 다이어그램을 그리기 위해서는 사용기관에서 필요한 데이터를 문장으로 기술해야 한다. 여기서 문제는 필요한 데이터들 문장을 문장으로 기술하는 데는 너무 많은 시간과 노력이 필요할 뿐만 아니라, 작성된 문장의 모호성도 문제가 될 수 있다.

교수 테이블

교수번호	교수이름	전공	직급	학과
8923	이휘소	핵이론	정교수	S01
7658	김진기	신학	부교수	G02
9453	박철수	반도체	부교수	E12

학과테이블

학과번호	학과이름	정원
S01	물리학	20
G02	신학과	25
E12	전자과	45

그림 1. 교수와 학과를 나타내는 테이블



그림 2. 교수와 학과관계의 ER 다이어그램

세 번째 이유로는 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[3, 7], 자료흐름도 작성이후 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(relation)는 ER 모델에서의 엔티티와 관계성(relationship)을 모두 나타낼 수 있다고 논문[5]에서 밝히고 있다. 따라서 데이터베이스 설계시 ER 모델을 사용하지 않고 관계형 테이블을 사용하더라도 스키마의 설계에는 문제가 발생하지 않는다.

III. 구조적 기법을 이용한 시스템 분석과 설계

본 논문에서 새로운 데이터베이스의 설계 방법을 설명하기 위한 예로 커피가공회사의 원료 구매 업무를 위한 데이터베이스 시스템을 설계한다. 구매업무의 최하위 계층도는 그림 3과 같이 작성되었다고 가정하고, 자료흐름도에 따라 다음과 같은 자료사전이 작성되었다고 하자.

- 커피원두재고목록 = 일자 + 커피원두 + 재고량
- 커피원두목록 = {커피원두}
- 커피원두 = 원두코드 + 원두이름 + 원두종류 + 원두단가
- 원두종류 = [레귤러 | 디카페인]
- 공급자화일 = {공급자}
- 공급자 = 공급자코드 + 공급자상호 + 공급담당자 + 공급자주소 + 공급자우편번호
- 구매신청화일 = {구매신청서}

구매신청서 = 구매신청번호+공급자
 +구매신청일자+수령예정일자 +커피원두
 +주문량+소계} +구매총액
 커피원두 인수목록
 = {커피원두+인수량+소계}
 인수내역화일 = {인수보고서}
 인수보고서 = 인수번호+구매신청번호+선적번호
 +인수담당자+인수일자+운송자+공급자 +
 {커피원두+인수량+상자수+소계}+인수총액
 지불요청내역 화일 = {지불요청서}
 지불요청서 =
 지불요청번호+구매신청번호+선적번호
 +인수번호+공급자+ 지불요청일자+
 {커피원두+인수량+소계} +총액+운송비
 +총비용

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

스키마를 생성한 다음 단계는, 각 테이블 사이의 관계를 파악하여, 테이블 사이의 참조무결성 제약 조건을 설정하고, 테이블 사이의 참여차수와 참여

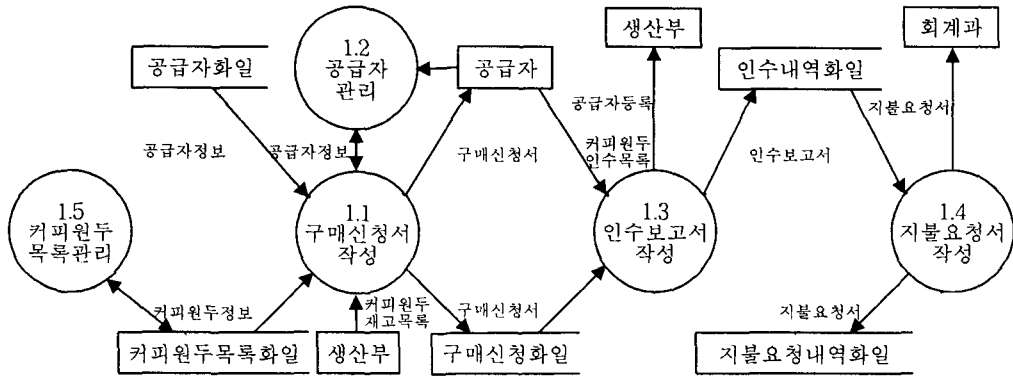


그림 3. 구매업무의 자료흐름도

IV. 자료 저장소를 이용한 관계형 데이터베이스 스키마의 설계와 테이블 관계도의 작성

관계합성[6]은 데이터베이스의 애트리뷰트를 먼저 추출하고, 애트리뷰트 사이의 함수적 종속성을 파악하여 관계형 데이터베이스를 설계하는 방법이다. 본 논문에서는 자료저장소를 구축하면서 파악된 애트리뷰트를 이용하여 관계합성 방법으로 데이터베이스 스키마를 설계한다. 자료저장소의 애트리뷰트를 기반으로 테이블을 구성한 후, 제 3정규형으로 정규화한 스키마는 다음과 같다.

비율을 알아보아야 한다. 이를 위해서 이미 도출된 8개의 테이블을 하나씩 차례로 비교하여 그림 4와 같은 테이블 관계도를 그린다. 테이블 관계도에 ER 다이어그램과 유사하나, 물리적 설계에서 참조무결성이나 데이터 갱신연산에 부과되는 제약을 설정하기 위해서 필요하다. 새로운 설계방법의 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 논

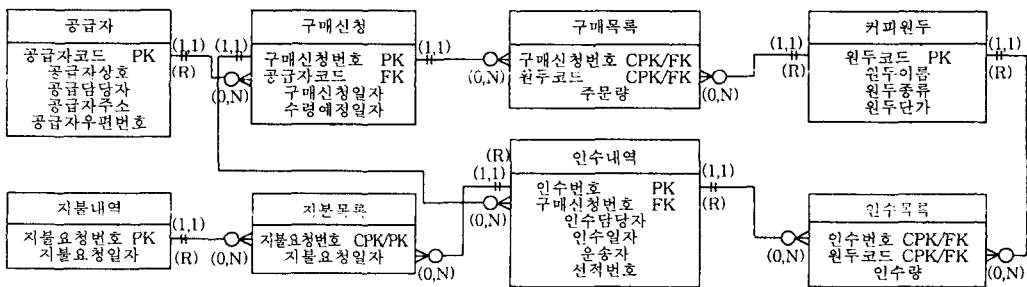


그림 4. 원두구매 업무를 위한 테이블 관계도

리적으로 작업을 진행할 수 있다.

V. 결 론

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 방법의 어려움과 ER 다이어그램 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 설계방법론에 의해서 단계적으로 관계형 데이터베이스 스키마를 설계할 수 있다. 먼저 구조적 설계방법을 적용하여 자료흐름도, 자료사전, 미니명세서를 작성한다. 자료저장소로부터 관계합성에 의해서 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 기능적인 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.

참고 문헌

- [1] 박재년, 구조적 시스템 분석과 설계, 정익사, 1992.
- [2] 우치수, Software 공학, 상조사, 1994.
- [3] Batini, Conceptual Database Design, Benjamin/Cummings, 1992.
- [4] Codd, E."A Relational Model for Large Shared Data Banks," CACM, 13/6, June 1970.
- [5] Date, C. "Entity/Relationship Modeling and the Relational Model." in C. Date and Hugh Darwen, Relational Database Writings 1989-1991, Mass.; Addison-Wesley, 1992.
- [6] Elmasri, R. & Navathe, S. Fundamentals of Database Systems, Addison-Wesley, 2000.
- [7] Kendall, P. Introduction to Systems Analysis & Design: A Structured Approach, IRWIN, 1994.