

의료정보 시스템과 데이터베이스

- 서울의대 의료정보학 강좌 -
2005.09.07

의료정보 데이터베이스

- Data and Database
- File System, Database System and Database Management System
- Historical overview and the Emergence of Relational Database
- Entity-Relationship Modeling
- Normalization
- DBMS and SQL (Structured Query Language)

데이터베이스의 활용

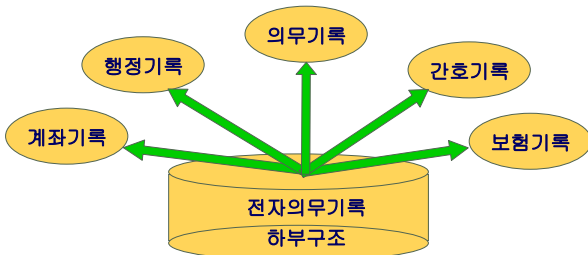
- 통신
 - 교통
 - 물류
 - 재정
 - 지식경영
 - 문헌정보
 - 사회사업
 - 국방
 - 교육
 - 과학연구
- ... 30년여의 짧은 연구역사
... 1000억불 시장의 형성

PsyBase 1.0

1995년 서울대학교병원 신경정신과에서 사용되기 시작한 국내 최초의 전자의무기록 PsyBase 1.0.

의무기록의 하부구조와 표면구조

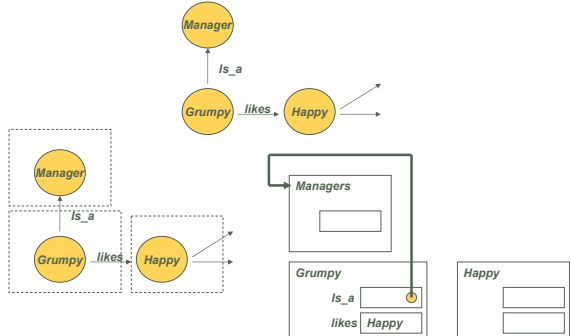
What is the unified view of records?



자료, 정보, 지식

- Data
- Information
- Knowledge

Data, Data Structure, Abstraction, Knowledge Representation, & Formalism



의료정보 데이터베이스

- Data and Database
- File System, Database System and Database Management System
- Historical overview and the Emergence of Relational Database
- Entity-Relationship Modeling
- Normalization
- DBMS and SQL (Structured Query Language)

파일 시스템의 문제점

- 중복 redundancy
- 일관성부재 inconsistency
- 모델링 불가능: 나이: -3
- 접근성 및 질의 인터페이스: access & query
- 파편화: fragmented, non-integrity
- 공유?
- 참조무결성: reference integrity, constraints
- 안전성: safety: system failure, access control



DBMS (Database Management System)

- Database is NOT DBMS
- Database system:
 - Schema: external, conceptual, internal
 - Database
 - Database Language:
 - DDL - definition
 - DCL - control
 - DML - manipulation
- DBMS

ORACLE

Microsoft SQL Server

새로운 데이터베이스: Challenge

- 방대한 데이터
- 복잡한 자료형
- 확장형 자료형
- 멀티미디어
- 논리연산 디비
- 분산형 데이터 처리
- 병렬형 데이터 처리

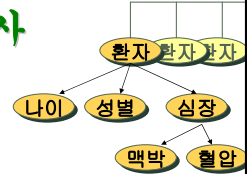
데이터베이스의 요구사항

- Efficiency: 자료 접근 및 갱신의 효율적 수행
- Resilience: 하드웨어 작동불량이나 소프트웨어 에러에도 불구하고 자료의 무결성을 손상하지 않고 견딜 수 있는 능력
- Access control: 복수 사용자의 동시 접속이 가능해야 하며 이러한 과정의 일관성이 유지되어야 한다. 사용자의 접근 위계질서 통제가 가능해야 한다.
- Persistence: 자료에 접근하는 응용프로그램의 동작 방식에 무관하게 장기적으로 자료의 무결성이 유지되어야 한다.

의료정보 데이터베이스

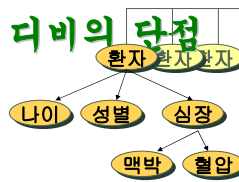
- Data and Database
- File System, Database System and Database Management System
- Historical overview and the Emergence of Relational Database
- Entity-Relationship Modeling
- Normalization
- DBMS and SQL (Structured Query Language)

데이터베이스 발달사



- Hierarchical (계층형):
root – nodes - leaves
- Network (네트워크형)
DAG (directed acyclic graph)
CODASYL(Conference on Data Systems Language)
- Relational (관계형)
- Object-oriented... (객체지향형...)

계층형 및 네트워크형 디비의 단점



- 특정한 질의에 대한 답을 구하기 위해서는 나무구조나 네트워크 구조를 향해가기 위한 매우 복잡한 프로그램을 작성해야 한다.
(예) 혈압이 140이상인 30세 이상 남자 환자는?
- 자료 구조를 변경해야 하는 경우(새로운 자료를 추가할 때마다 생기는 일로서) 응용프로그램을 그에 맞추어 수정해야 한다.
즉, 자료 조작(접근, 검색, 갱신 등)이 자료구조에 종속
→ 과다한 데이터 관리비

관계형 데이터베이스의 시작

- E. F. Codd in 1970-72
- Tabular data structure (relations)
- High level non-procedural(declarative) query language (predicates)
- Query optimizer

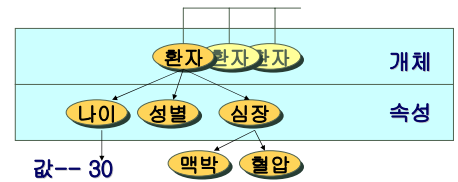
즉, 자료구조에 독립적인 자료처리 (질의) 언어체계
"Abstraction Barrier"

→ 프로그래머 생산성 향상 및 최종사용자의 직접적인 자료접근 가능

의료정보 데이터베이스

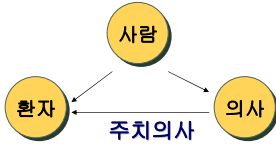
- Data and Database
- File System, Database System and Database Management System
- Historical overview and the Emergence of Relational Database
- Entity-Relationship Modeling
- Normalization
- DBMS and SQL (Structured Query Language)

개체(Entity)와 속성(Attribute)



- 개체: 정보의 기본 단위
- 속성: 개체를 기술하는 특성
- 값: 속성의 값

관계(Relationship)

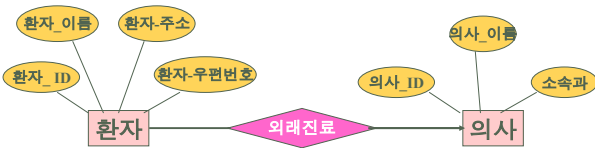


ER Modeling 개체(Entity) - 관계(Relationship)

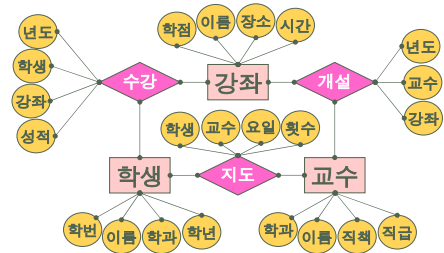
- 개념적 데이터 모델링
- 개체와 개체 사이의 관계를 E-R diagram으로 표현
- E-R Diagram의 기본요소
- Relational Model로 쉽게 연결됨



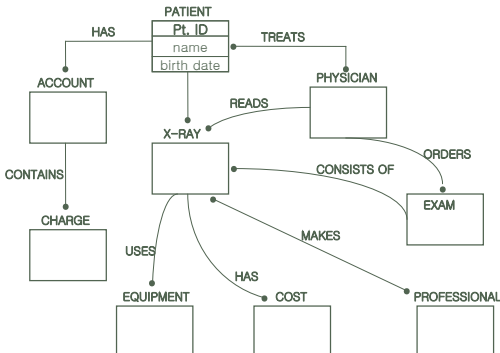
Entity Relationship Modeling



Entity Relationship Modeling



Entity Relationship Diagram



관계(레이블, Relation)와 레코드

고유번호	의사성명	환자병력번호	환자명

Relation (or Table)
 Tuple (or Record or Instance)
 Attribute

주키와 보조키 (Primary & Secondary keys)

- Identifier
- Primary key: uniquely identify
- Secondary key: identify (ex. Foreign key)

고유번호	의사성명	병동	환자번호	환자명
123456	홍길동	12	2432554325	김이순
432521	이몽룡	12	4331215955	황단이
532665	초롱이	13	5385358859	한림대
123456	홍길동	13	1243255432	동아리
435452				

의료정보 데이터베이스

- Data and Database
- File System, Database System and Database Management System
- Historical overview and the Emergence of Relational Database
- Entity-Relationship Modeling
- Normalization
- DBMS and SQL (Structured Query Language)

관계형 데이터베이스의 예

What's wrong with this table

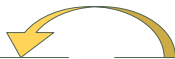
고유번호	의사성명	병동	환자번호	환자명
123456	홍길동	12	2432554325	김이순
432521	이몽룡	12	4331215955	황단이
532665	초롱이	13	5385358859	한림대
123456	홍길동	13	12432554325	동아리
435452				

관계형 데이터베이스의 예



고유번호	의사성명	담당의	병동	환자번호
123456	홍길동	123456	12	132114324
432521	이몽룡	432521	12	559435255
532665	초롱이	532665	13	582860436
435452		123456	14	998140944
		435452		

관계형 데이터베이스의 예



담당의	병동	환자번호	환자번호	환자명	나이	성별	주소
123456	12	132114324	132114324	홍길동	30	m	
432521	12	559435255	559435255	이몽룡	19	m	
532665	13	582860436	582860436	초롱이	22	f	
123456	14	998140944	998140944				
435452							

스키마 (Schema)

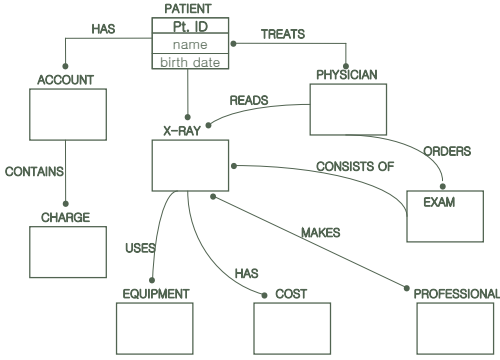
Doctor(*Doctor_id*, *Doctor_name*, *Doctor_address*, *DateOfBirth*, *Department*)

Patient(*Patient_id*, *Patient_name*, *Patient_address*, *Patient_phone*,

DateOfBirth, *Sex*)

Doctor_Patient(*Doctor_id*, *Patient_id*)

Entity Relationship Diagram



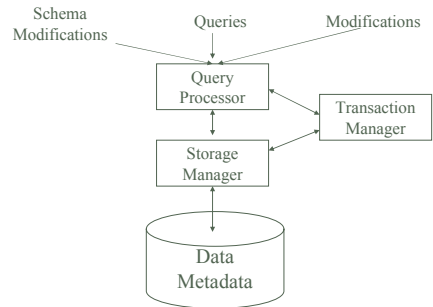
데이터베이스 정규화

- **1NF**
 - **2NF**
 - **3NF**
 - **4NF**
 - **5NF**
- **개체무결성(entity integrity):** 기본키는 각 레코드의 기본이 되므로 null 값을 가질 수 없다.
 - **참조무결성(referential integrity):** 외래키는 참조되는 테이블에 존재하는 값이어야 한다.
 - **도메인무결성(domain integrity):** 속성값은 해당 속성이 소유할 수 있는 속성값을 가져야 한다. (예: $0 < \text{나이} < 200$)
 - **사용자정의 무결성(general integrity):** 의료정보학 학점은 A, B, C, D, F로...?

의료정보 데이터베이스

- *Data and Database*
- *File System, Database System and Database Management System*
- *Historical overview and the Emergence of Relational Database*
- *Entity-Relationship Modeling*
- *Normalization*
- *DBMS and SQL (Structured Query Language)*

DBMS Overview



구조화 질의 언어 SQL: Structured Query Language

- **Relational algebra** 관계대수
- **Relational calculus** 관계함수
- **Select** → 행 (R WHERE Condition)
- **Project** → 열 (R [A1, A2])
- **Join** → 두 R의 결합 (R JOIN(A1 = B1) S)

구조화 질의 언어 SQL: Structured Query Language

- 한 질의의 결과도 역시 관계(Relation; table)이다
- 그러므로 closer가 생기고, 즉, 연속적인 연산이 가능하다.
- 이러한 질의가 자료구조에 무관하다.
- 자료의 참조무결성을 유지할 수 있다.

구조화 질의 언어 SQL: Structured Query Language

- 의사(면허번호, 성명, 소속과, 생년월일, 성별)
- 담당의사(의사번호, 병력번호, 환자명)
- {의사 JOIN(면허번호 = 의사번호) 담당의사}[성명, 병력번호]
- **SELECT** 성명, 병력번호
FROM 의사, 담당의사
WHERE 의사.면허번호 = 담당의사.의사번호

트랜잭션 (Transaction)

- “a sequence of operations” that must appear 'atomic' when executed
- otherwise, *inconsistent* database state results
- **Concurrency control**: 각 트랜잭션을 독립실행
- **Recovery**: 각 트랜잭션은 all or none
serializability & schedulers

관계형 데이터베이스의 발달

- High-level relational query languages
- Query optimization:
→ high-level language into planed series.
- Normalization Theory
→ eliminating redundancy and certain logical anomalies
- Pages → to minimized the cost of access
- Buffer management algorithms
- Indexing techniques

분산데이터베이스 Distributed

- Business is being more and more distributed
- Distributed vs. fragmented

W3EMRS: The 1st Generation

