

데이터 모델 링

백앤드강의

데이터모델링 기법

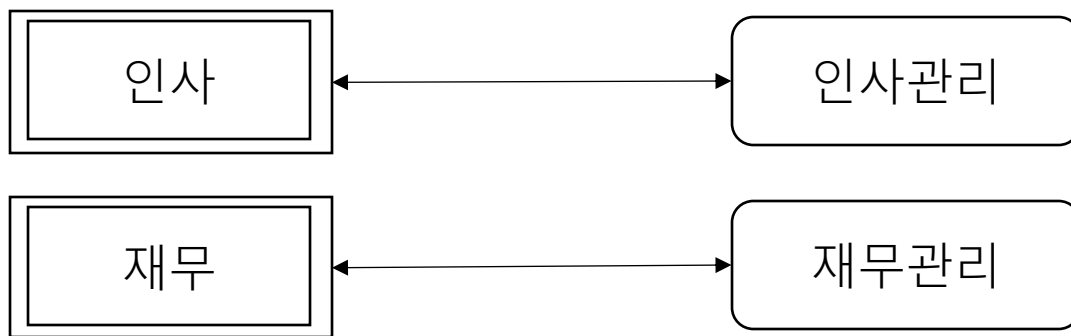
- 주제영역
- 엔티티 타입
- 관계
- 속성
- 식별자
- 서브타입
- 정규화

주제영역 정의

- 업무에서 관리하고자 하는 데이터의 그룹
- 데이터 세분화의 기초 – 하향식 접근
- 하위에 주제영역 또는 엔티티 타입을 포함
- 일반적으로 관련된 업무 기능이 존재
- 정보전략계획수립(ISP) 단계에서 분석 수행

주제영역 명명

- 현업에서 보편적으로 사용하는 업무 용어 부여 : 인사, 생산, 판매, 구매, 재무 등
- 주제영역에 대해 유일한 단수명사형 사용
- 중심 엔티티(주제영역 대표 엔티티) 타입명과 동일할 수 있음
- 데이터의 그룹을 의미하는 이름 부여 : 업무활동을 의미하는 이름 배제



주제영역 계층

- 최상위 주제영역(Root Subject Area)
 - 분석 대상 영역의 전체 범위 → 삼성생명보험주식회사 또는 보험계약
 - 단위 주제영역 (Primitive Subject Area)
 - 더 이상 주제영역을 포함하지 않고, 엔티티 타입과 관계로 구성된 최하위 주제영역
 - 중심 엔티티가 존재
 - 단위 업무기능(Primitive Function)과 대응
 - 단위업무 기능 : 업무기능의 최하위 기능

주제영역 1차 수준

최상위 주제영역

삼성생명주식회사

보험 당사자

상품

계약

지급

주제영역 2차 수준

최상위 주제영역

삼성생명주식회사

보험 당사자

계약

지급

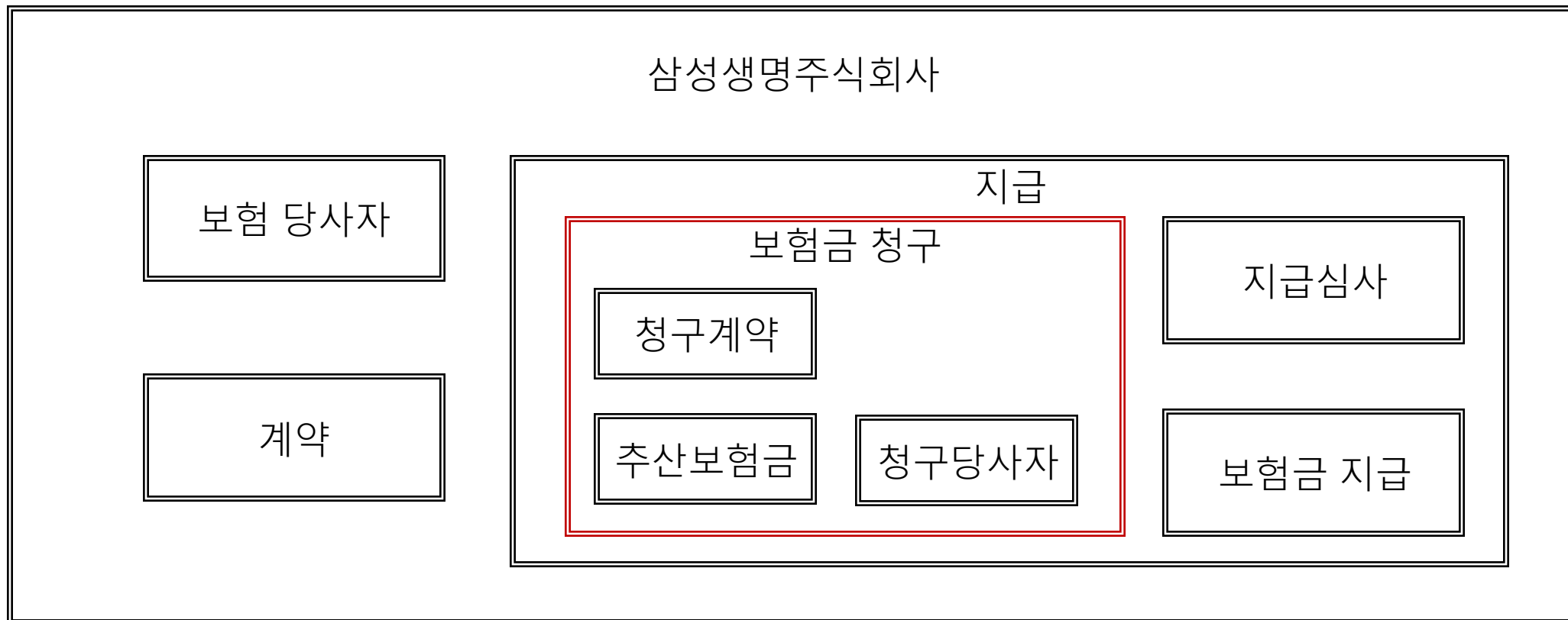
보험금 청구

지급심사

보험금 지급

단위 주제영역

최상위 주제영역



중심 엔티티 타입

최상위 주제영역

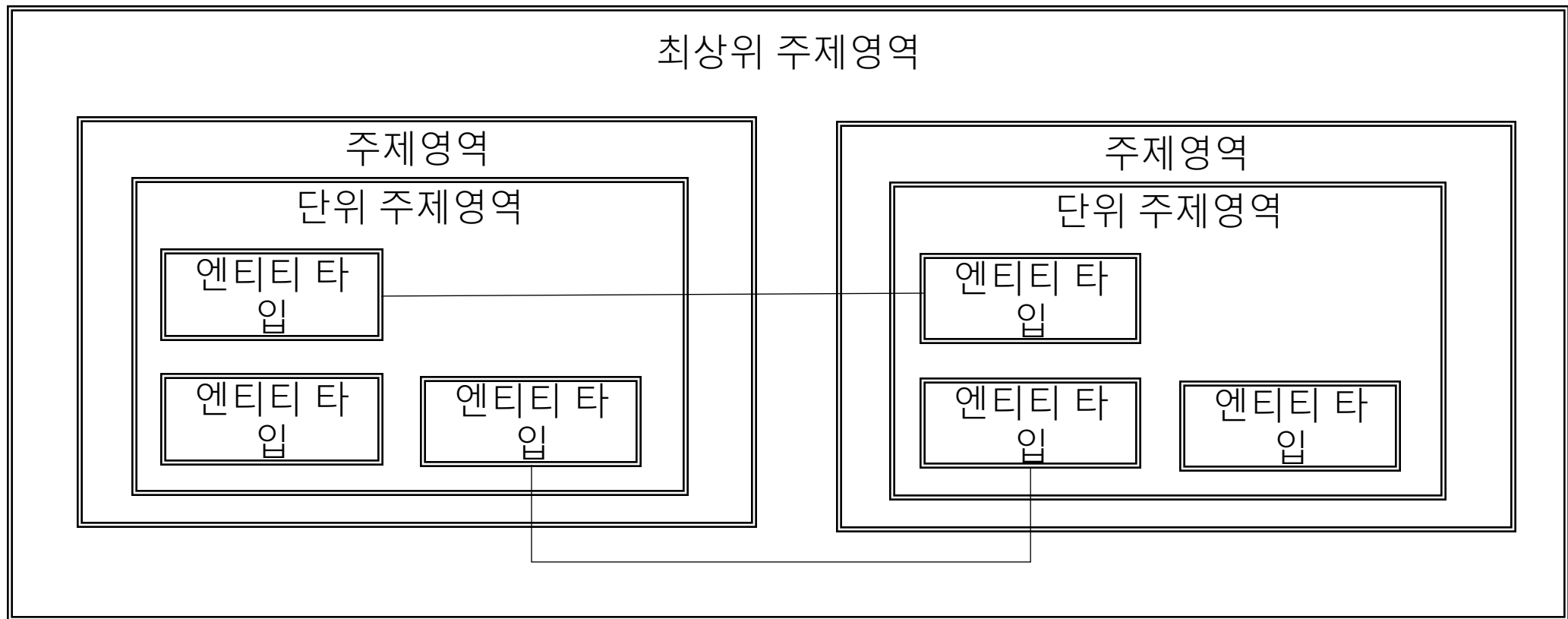
삼성생명주식회사

보험 당사자

보험당사자
주소

주제영역 표현

최상위 주제영역





주제영역 활용

1. 목적

- 데이터의 계층적 구조 파악
- 업무기능(function)과 병행하여 분석 - 품질 확보

2. 장점

- 데이터 및 업무활동의 품질 보증 모델
- 프로젝트 관리 용이
- 모델 개발 조정 용이

주제영역 도출

- 업무에서 사용되는 데이터의 명사형 도출
- 업무기능의 이름으로부터 도출
- 하향식 접근 방법
- 상향식 접근 방법
- 분석단계에서 도출

주제영역 정의내용

레벨	주제영역명	설명	포함 엔티티 타입
1레벨	지급	보험금 지급	
2레벨	보험금청구	지급 > 보험금 청구 업무	
3레벨	청구관련계약	보험금 청구 업무 중 청구관련 계약	청구계약, 계약상태...

엔티티타입 정의(Definition)

- 업무에서 관리하고자 하는 데이터의 한 형태
 - 예) 사람, 사물, 장소, 사건등
- 단위주제영역에 포함
- 업무프로세스와 대응
- 정보전략수립계획(ISP)단계에서 시작, 업무 분석(BA)단계에서 분석완료
- 데이터베이스에서 테이블로 구현
- 표현기호

엔티티타입명

엔티티 타입의 분류(Classification)

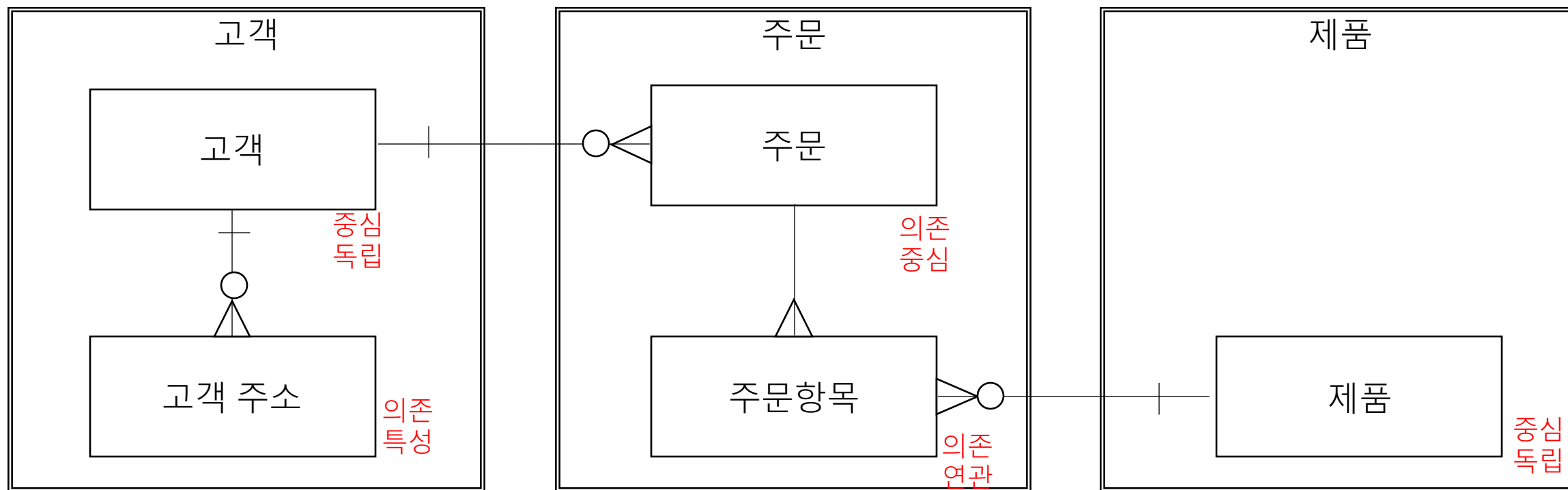
- 일반적인 분류
 - 실체(Tangible) 엔티티 타입
 - 비 실체(Intangible)엔티티 타입
 - 개념(Conceptual) 엔티티 타입
 - 사건(Event) 엔티티 타입
- 모델관점의 분류
 - 중심(Central) 엔티티 타입
 - 독립(Independent) 엔티티 타입
 - 의존(Dependent) 엔티티 타입
 - 특성(Characteristic) 엔티티 타입
 - 연관(Associative) 엔티티 타입

일반적인 분류

- 실체 엔티티 타입
 - 물리적인 형태가 있는 엔티티 타입
 - 안정적이며, 지속적으로 활용되고 공유의 대상인 경유가 많음
 - 비교적 도출이 쉬우며, 우선 도출 대상 (직원, 제품, 고객, 과목 등)
- 개념 엔티티 타입
 - 물리적인 형태가 없으나 관리의 필요에 의해 사용되는 엔티티 타입(조직, 계정, 상품)
- 사건 엔티티 타입
 - 업무를 수행하는 행위와 관련된 엔티티 타입 (발송, 주문, 구매, 계약, 청구 등)

모델관점의 분류

- 중심, 독립, 의존, 특성, 연관 엔티티 타입



모델관점의 분류

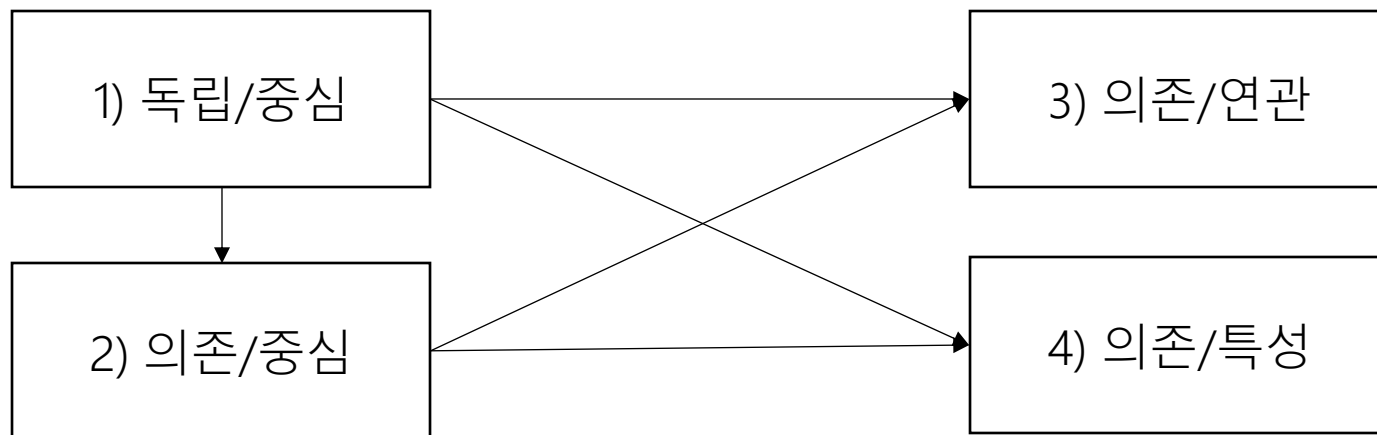
- 중심 엔티티 타입
 - 단위 주제영역에서 가장 핵심이 되는 엔티티 타입
 - 일반적으로 단위 주제영역과 동일한 이름 (고객, 주문, 제품 등)
- 독립 엔티티 타입
 - 다른 엔티티 타입의 존재와는 독립적으로 존재가 가능한 엔티티 타입
 - 일반적으로 중심 엔티티 타입이면서 공유의 대상이 됨 (고객, 제품)
- 의존 엔티티 타입
 - 다른 엔티티 타입에 의존하는 엔티티 타입(고객주소, 주문, 주문항목)

모델관점의 분류

- 특성 엔티티 타입
 - 특정 엔티티 타입의 특징을 상세히 설명해주는 엔티티 타입
 - 항상 의존 엔티티 타입이 된다 (고객 주소, 고객 전화번호 등)
- 연관 엔티티 타입
 - 두개 이상의 엔티티 타입에 의존적인 엔티티 타입
 - 다대다(M:N)의 관계를 해결하면서 나타나는 엔티티 타입
 - 항상 의존 엔티티 타입(주문 항목)

모델관점의 분류

- 하향식으로 모델 전개
 - 독립/중심, 의존/중심 엔티티 타입으로 부터 출발
- 모델의 검토 및 품질보증
 - 분석의 상세화 수준에 따라 파악



엔티티와 엔티티 타입

- 엔티티 타입 : 공통적인 정의(속성)을 가지는 모든 엔티티 집합
- 엔티티(Entity)
 - 기업에서 관리하는 실제 데이터의 건
 - 엔티티 타입의 개별 발생건

사번	성명	주소	전화번호	...
123456	박길동	서울시 양천구	111-2222	...
678901	홍길동	서울시 금천구	222-3333	...

엔티티 타입 명명

- 현업에서 사용하는 업무용어 부여
 - 현업과 조정하여 업무용어 및 엔티티 타입명 표준화
- 엔티티 타입에 대해 유일한 단수명사형 사용
 - 단위주제영역명과 동일
 - 약어 사용배제
- 별명은 별도로 관리
 - 동일한 엔티티 타입에 대해 업무적으로 여러가지 이름을 사용해야 할 경우 대표적 이름 부여하고, 별칭은 별도로 관리
 - 회원 (고객, 직원, 등등)

엔티티 타입 도출

- 업무에서 사용하는 데이터의 명사형 도출
- 아키텍처 모델로부터
- 모델 관점의 분류 활용
- 일반적인 분류 활용
- 업무 프로세스 이름으로 도출

엔티티 타입 정의내용

- 엔티티 타입 명
- 엔티티 타입 설명
- 엔티티 타입 특성(예상 건수, 예상 증가율 등)
- 포함되는 단위 주제영역 이름
- 별명

엔티티 타입 특성

- 엔티티의 예상건수
 - 보관주기 동안 발생할 수 있는 엔티티 건수
 - 과거의 기록 또는 향후 업무 예측을 통해 산정
 - 최소/평균/최대값으로 표현
 - 데이터베이스 구축 시 용량산정의 기초자료로 활용
- 예상 증가율
 - 보관주기 동안 엔티티 발생건수의 증가/감소 비율
- 영문명
 - 테이블 명, 명명 규칙에 따른 영문명 사용(INSACCTB 등)

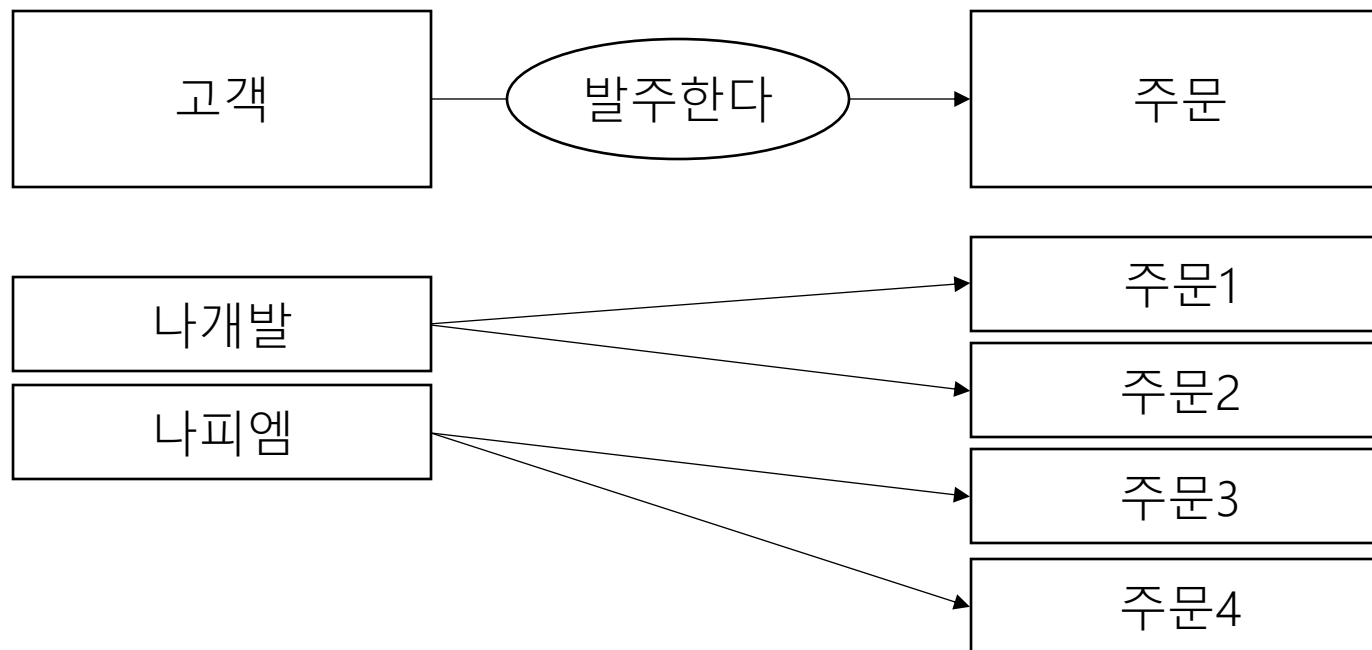
관계 정의(Definition)

- 엔티티 타입 간에 관련이 생기게 되는 업무적 이유
- 데이터의 양방향 업무규칙을 표현
- 관계를 통해 정보 활용 가치 상승
- 외래키(Foreign key)로 구현되어 참조 무결성으로 데이터 정합성 유지



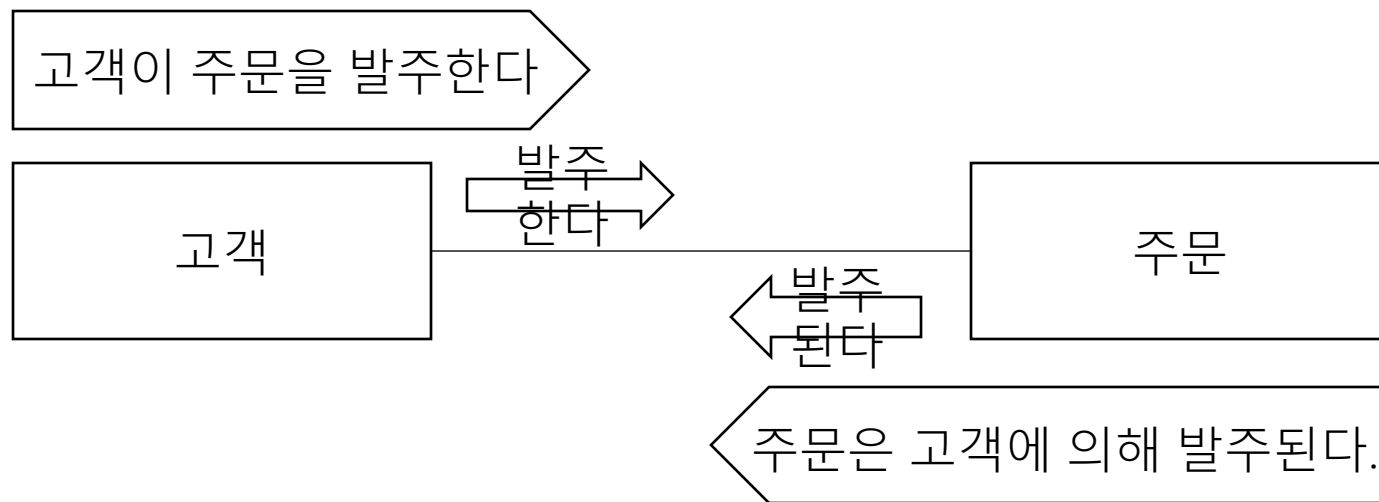
페어링과 관계

- 관계: 두개의 엔티티 타입간의 연관성
- 페어링 : 관계의 개별 발생 건



관계 멤버십

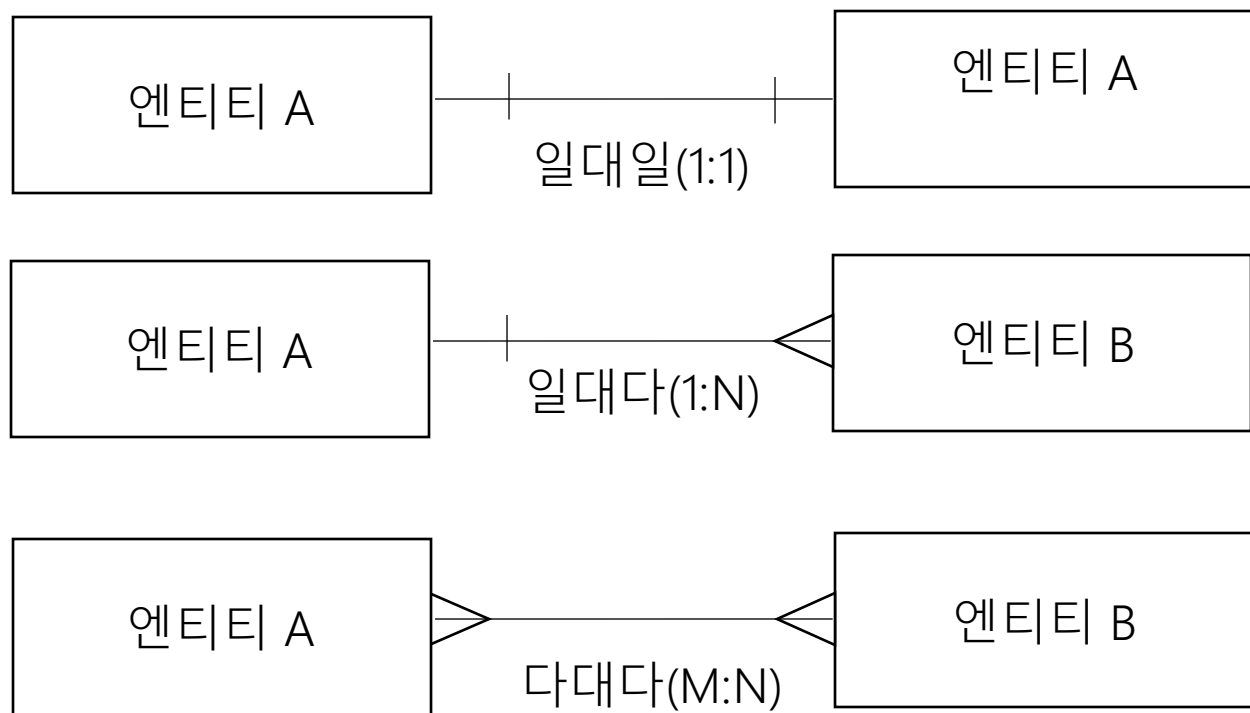
- 관계를 양쪽 방향에서 각각 바라본 관점
- 하나의 관계는 항상 두 멤버십으로 구성
 - 멤버십 주체는 하나의 특정 엔티티
 - 두 멤버십에 각각의 업무 규칙 발생



관계멤버십의 표현요소

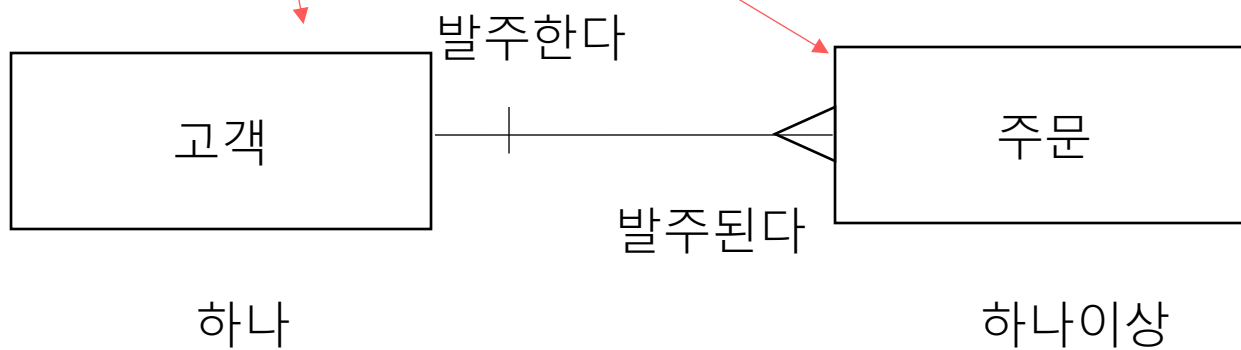
- 카디널리티(Cardinality)
 - 의미 : 하나의 엔티티가 가질 수 있는 페어링 수
 - 유형 : 일대일(1:1), 일대다(1:N), 다대다(M:N)
- 선택성(Optionality)
 - 의미 : 페어링 존재여부
 - 유형 : 필수적, 선택적
- 이동가능성(Transferability)
 - 의미 : 동일한 관계에 대해서 페어링의 변경 가능 여부
 - 유형 : 고정, 이동가능 등

카디널리티 유형

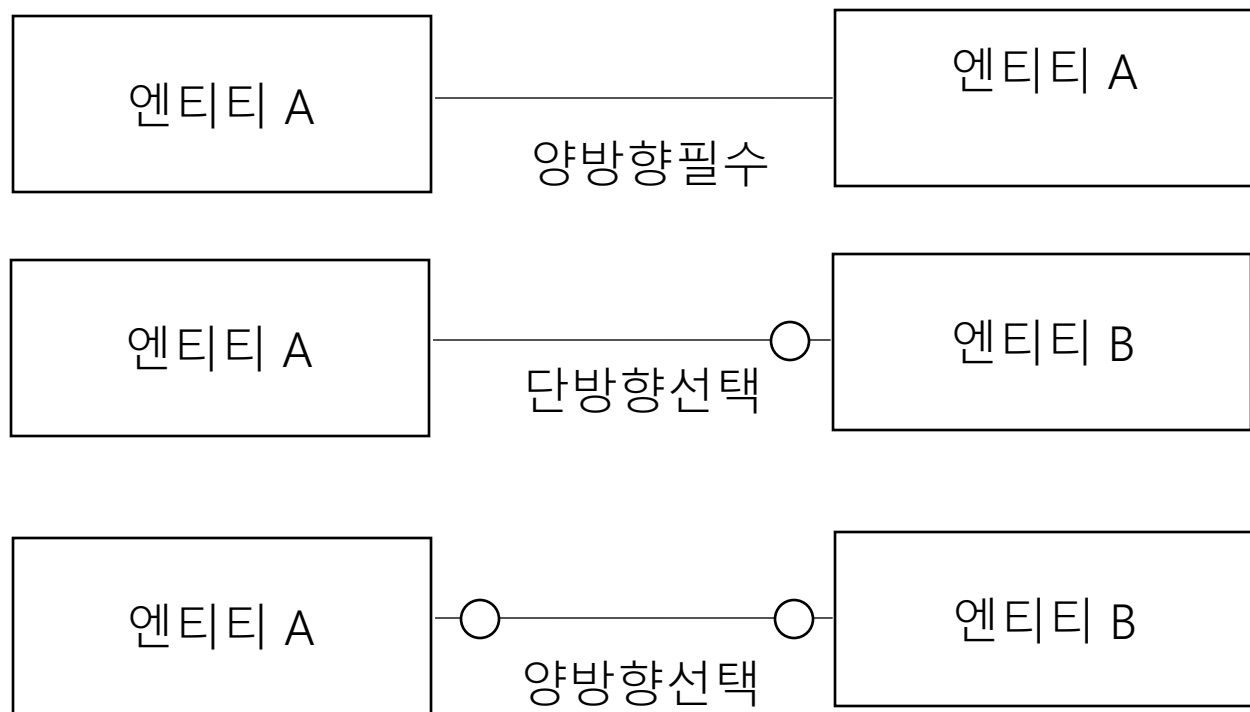


카디널리티의 표현

- 각 고객은 하나의 **하나 이상의** 주문을 발주한다.
- 각 주문은 하나의 **하나의** 고객에 의해 발주된다.

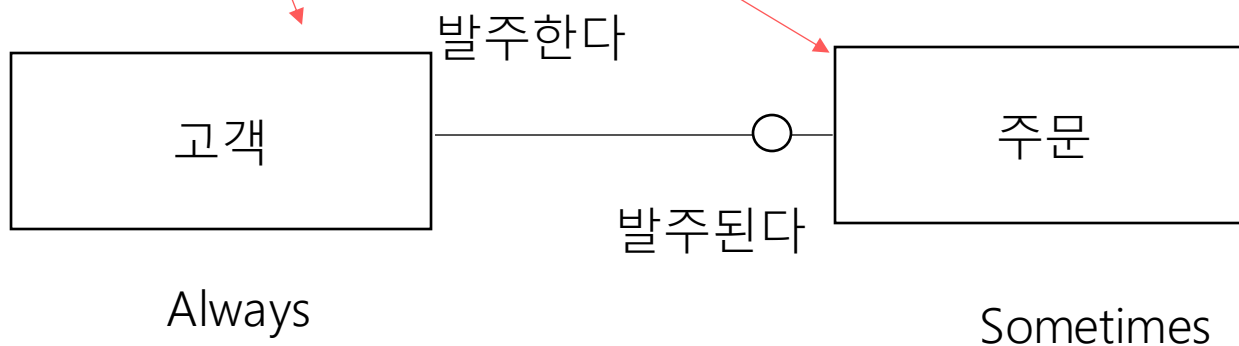


선택성의 유형



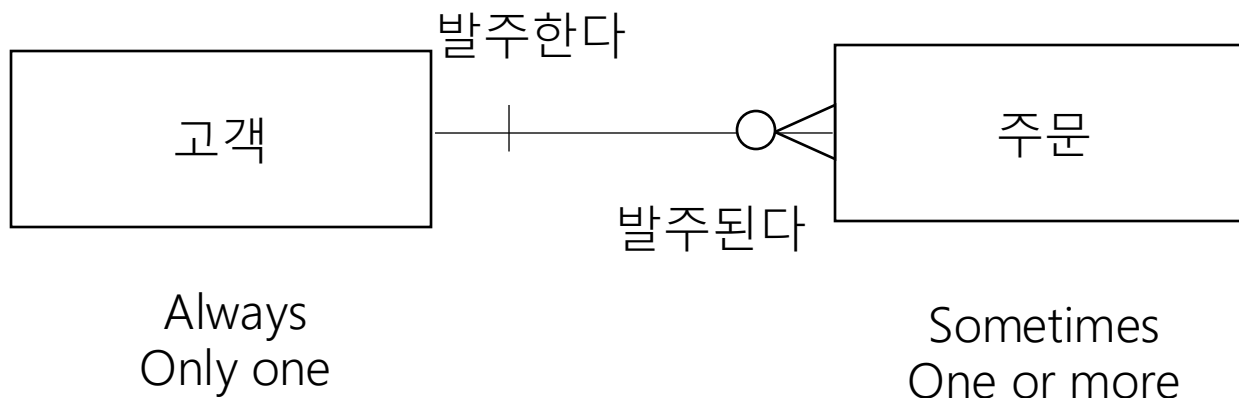
선택성의 표현

- 각 고객은
 - 각 주문은
- 항상
때때로
- 항상
때때로
- 주문을 발주한다.
- 고객에 의해 발주된다.



카디널리티-선택성 결합

- 각 고객은 때때로 하나 이상의 주문을 발주한다.
- 각 주문은 항상 하나의 고객에 의해 발주된다.

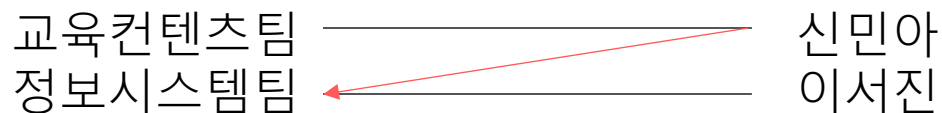


이동 가능성(Transferability)

- 고정(Fixed) : 다른 페어링으로 변경 불가

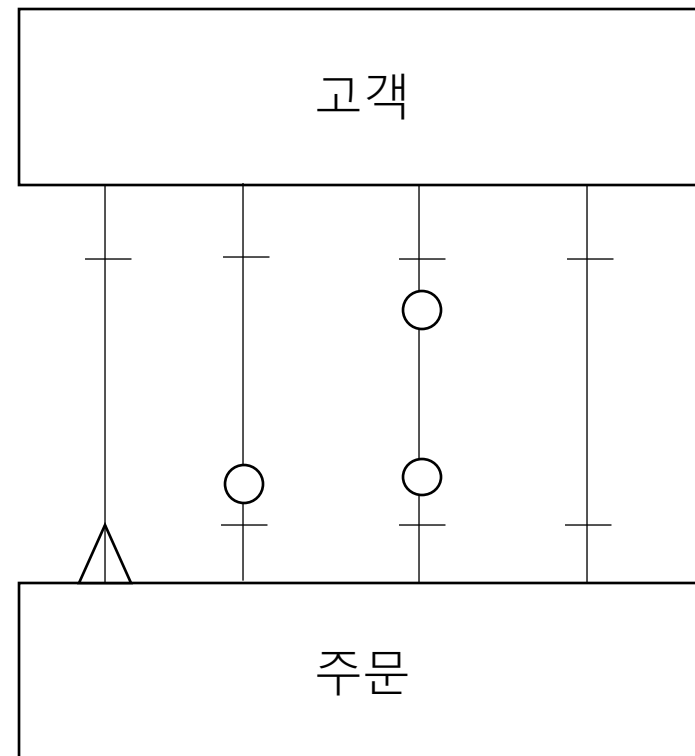


- 이동 가능: 다른 페어링으로 변경 가능



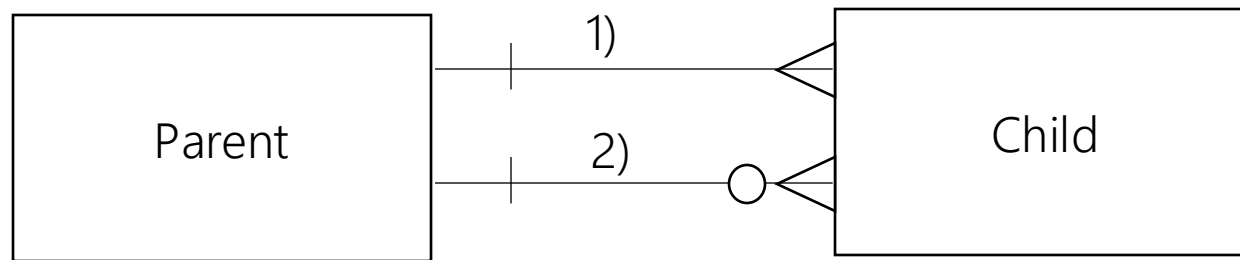
관계에 따른 엔티티타입의 계층

- 상위(Parent) 엔티티 타입
 - 카디널리티 고려
 - 선택성 고려
- 하위(Child) 엔티티 타입
 - 상위 엔티티 타입의 반대로 작용
 - 외래키(Foreign Key) 발생



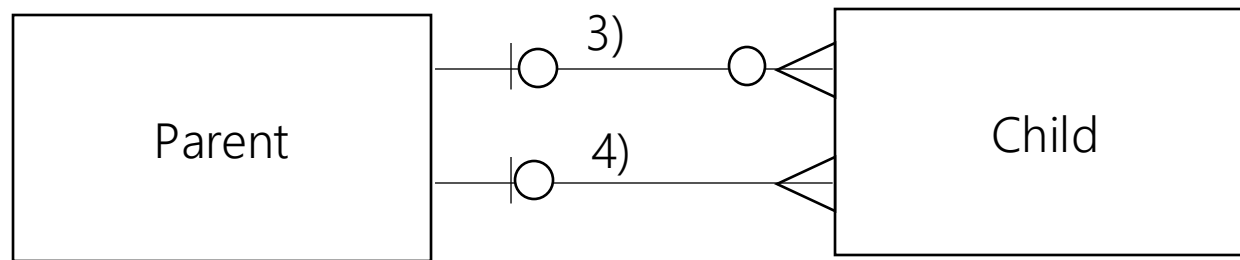
선택적 관계

- 1)의 경우 (Parent 필수/Child필수)
 - 상위 엔티티 생성시점에 하위 엔티티와 페어링 반드시 발생 (1 or M)
 - 하위 엔티티 생성시점에 상위 엔티티와 페어링 반드시 발생 (only 1)
- 2)의 경우 (Parent 필수/Child선택)
 - 상위 엔티티 생성시점에 하위 엔티티와 페어링 선택적 발생 (0, 1, M)
 - 하위 엔티티 생성시점에 상위 엔티티 페어링 반드시 발생 (only 1)



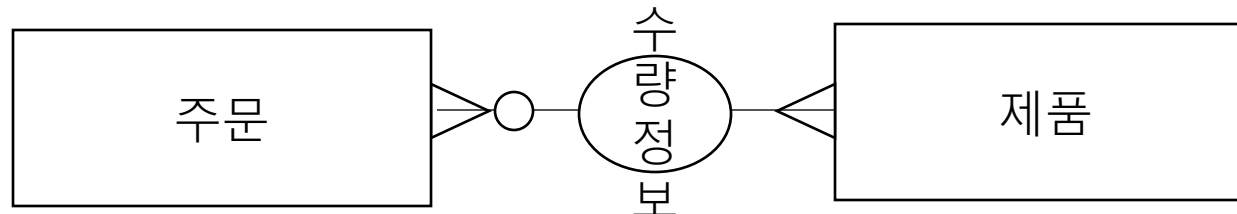
선택적 관계

- 3)의 경우 (Parent 선택/Child선택)
 - 상위 엔티티 생성시점에 하위 엔티티와 페어링 선택적 발생 (0,1,M)
 - 하위 엔티티 생성시점에 상위 엔티티와 페어링 선택적 발생 (0 or 1)
- 4)의 경우 (Parent 선택/Child필수)
 - 상위 엔티티 생성시점에 하위 엔티티와 페어링 반드시 발생 (1 or M)
 - 하위 엔티티 생성시점에 상위 엔티티 페어링 선택적 발생 (0 or 1)



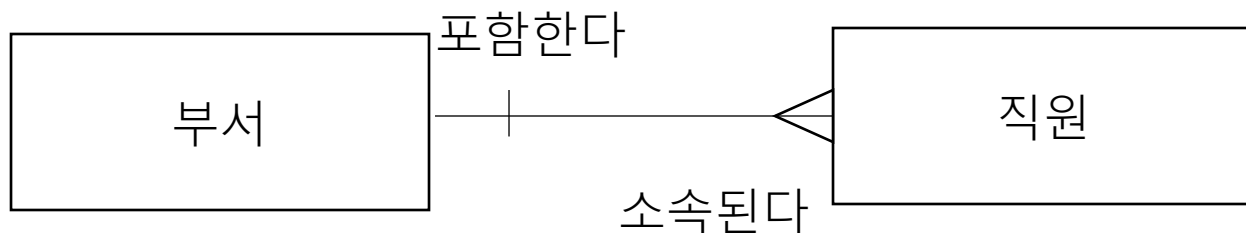
다대다(M:N)의 해결

- RDB에서 물리적 구현 불가능
- 해결방법
 - 관계에 숨어있는 추가적인 숨은 정보 파악
 - 엔티티타입을 추가하여 다대다(M:N)의 관계를 일대다(1:M)로 해결
 - 추가된 엔티티타입에 일대다(1:M)의 관계(선택성 상속) 반영
 - 추가된 엔티티타입에 속성 반영



관계명명

- 두개의 관계멤버십에 각각 부여
- 현업에서 사용하는 간결한 동사형 표현
 - 두 엔티티 타입간의 업무적 연관성을 나타내는 이름 부여
 - 현재시제를 사용
- 업무적 의미가 없거나 애매모호한 용어 배제



관계도출

- 업무에서 사용하는 동사를 통해서 도출
 - 업무규칙에 나타나는 동사 찾기
 - 두 엔티티 타입을 연결하는 동사 확인
 - 업무에서 필요한 정보가 둘 이상의 엔티티 타입을 조합
 - 화면, 보고서에 나타나는 엔티티 타입 존재 확인
- 아키텍처 모델로부터 출발
- 모델 관점에서 분류 활용
 - 독립/중심 엔티티 타입간 관계 존재여부 확인 등

관계 정의 내용

- 관계명
- 관계의 카디널리티/선택성
- 설명
- 특성
 - 카디널리티 개수
 - 선택성 비율
 - 이동가능성 여부
- 참조 무결성 규칙

속성

- 정보를 구성하는 최소단위 데이터
- 엔티티 타입의 특징을 표현
- 업무 분석(BA)단계에서 시작해서, 설계단계에서 분석완료
- 데이터베이스에서는 컬럼으로 구현

사번	성명	주소	전화번호	
1111	이지아	서울시 양천구 신정동	010	
2222	김효진	서울시 강남구 서초동	111	

속성값

- 각 엔티티들이 실제로 갖는 항목 값
- 각 속성은 특정 시점에 하나의 값만 존재
 - 시간이 경과하면서 변경 가능
- 예 : 사원 (엔티티 타입)

사번	성명	주소	전화번호	
1111	이지아	서울시 양천구 신정동	010	
2222	김효진	서울시 강남구 서초동	111	

속성 명명

- 현업에서 사용하는 업무 용어 부여
- 엔티티 타입내에서 유일한 단수 명사형 사용
 - 이름만으로 의미와 내용을 연상
 - 의미의 명확성을 위해 수식어 사용
 - 약어 사용배제
- 별명(Alias)는 별도로 관리
 - 동일한 속성에 대해 업무적으로 여러가지 이름을 사용해야 하는 경우에는 대표적 이름 부여, 별명은 별도 기록

속성 도출

- 업무에서 필요로 하는 데이터 항목 도출
 - 엔티티 타입으로부터 도출
 - 정보수집 소스 활용
- 정보수집 소스
 - 파일 레이아웃
 - 화면 UI 등 데이터 항목 참고
- 속성 타입은 분석과정, 설계 개발 과정에서 지속 도출

속성 정의 내용

- 속성 명
- 속성 설명
- 속성 특성(도메인, 길이, 유형, 선택성, 영문명)
- 속성값
- 별명

속성 설명

- 속성을 이해하는데 도움을 줄 수 있도록 업무적으로 기술
 - 코드성 속성의 경우 코드값 정의
 - 추출형 속성은 계산식 정의
 - 필요시 무결성 조건 정의
- 속성 설명의 활용
 - 설계 및 개발자에게 속성에 대한 업무적 이해
 - 데이터 사전 작성 시 활용
 - 도움말 시스템 구축 시 활용

도메인

- 속성에 대한 성질을 그룹핑한 개념
 - 항목 표준화에 따르는 사항으로 데이터관리자가 설정
 - 모든 업무영역에서 같은 도메인 집합을 사용
- 도메인의 예

도메인	속성 예
숫자	입금액, 구매수량, 이자, 점수
문자	사원이름, 제품명, 주소, 고객명
일자	입사일자, 생년월일, 계약일자
시간	거래시간, 마감시간,

속성 길이

- 현재 업무에서 사용 중인 최대 길이 지정
- 향후 변경요건을 고려해서 지정
 - 예 : 일자의 경우 안정성을 고려해서 8자리로 지정
- 숫자 도메인의 경우 소수점 이하 길이 지정
- 속성 길이는 향후 데이터베이스 필드 길이로 구현

속성 유형

- 기본형
 - 엔티티 타입을 기본적으로 표현하는 데이터
 - 업무로부터 직접 도출(예: 사원명)
- 추출형
 - 계산을 통해 도출되는 데이터(예:주문 총액)
- 설계형
 - 업무의 운영을 단순화 하기 위해 사용(예: 제품코드)

속성의 선택성

- 판단기준
 - 엔티티가 생성되는 시점에 속성값이 존재해야 하는지의 여부
 - 업무 규칙에 따라 결정
- 유형
 - 필수 : 특정 속성 값이 반드시 존재
 - 선택 : 특정 속성 값이 존재하지 않아도 됨

사번	성명	주소	차량번호	결혼기념일
1111	이지아	서울시 양천구 신정동	041111	
2222	김효진	서울시 강남구 서초동	051111	

속성값 지정

- 기본값
 - 입력시점에 값이 정의되지 않았을 경우 기본값 지정
 - 지정된 값 또는 범위에 의한 값 설정(예: 입력일, 주문총액)
- 허용값
 - 입력이 허용되는 속성값 지정(검증)
 - 지정된 값 또는 범위에 의한 값 설정

복합 속성

- 단일 값으로 취급 가능한 속성값의 집합으로, 대표명을 갖는 속성
 - 구성요소가 각각의 독립적인 의미를 가지는 경우
 - 별도의 관리목적이 있을 경우
- 복합 속성의 예

계좌번호	150	74	23	123456
	지점번호	상품종류	개설연도	일련번호

코드

- 활용하고자 하는 데이터를 약어 또는 기호로 함축하여 사용
 - 데이터의 길이를 단축시켜 데이터의 활용을 용이하게 함
 - 고객등급: (A, B,C, D...)
- 코드의 적용유형
 - 업무적인 필요에 의해 사용
 - 전산화 과정에서 사용

코드 속성

- 코드성 속성에 대해 코드 값을 할당하고 의미정의
- 코드 값은 향후 확장성 고려
- 코드 값이 여러 업무에서 사용되는 경우 통합코드로 일관성 유지
- 코드 목록 및 명세 작성
 - 코드 목록 : 분석 대상 업무영역에서 사용되는 전체 코드 목록
 - 코드 명세 : 각 코드별 값과 그 의미를 기술

식별자 정의

- 엔티티 타입의 각 엔티티를 유일하게 식별해 주는 구분자
 - 속성, 관계 또는 이들의 조합으로 구성
 - 유일성을 보장하는 최소한의 속성 또는 관계로 구성
 - 테이블의 유니크 인덱스로 구현(Primary Key)
- 식별자를 정의한 후 식별자의 구성요소를 정의
 - 단일 식별자 : 하나의 구성요소로만 정의
 - 복합식별자 : 다수의 구성요소로 구성
- 예 : **직원번호**, 직원성명, 주민번호, 직원주소, 담당업무

식별자의 구성요소

- 하나 혹은 다수의 속성으로 구성
- 하나 또는 다수의 관계로 구성
- 속성과 관계의 조합으로 구성



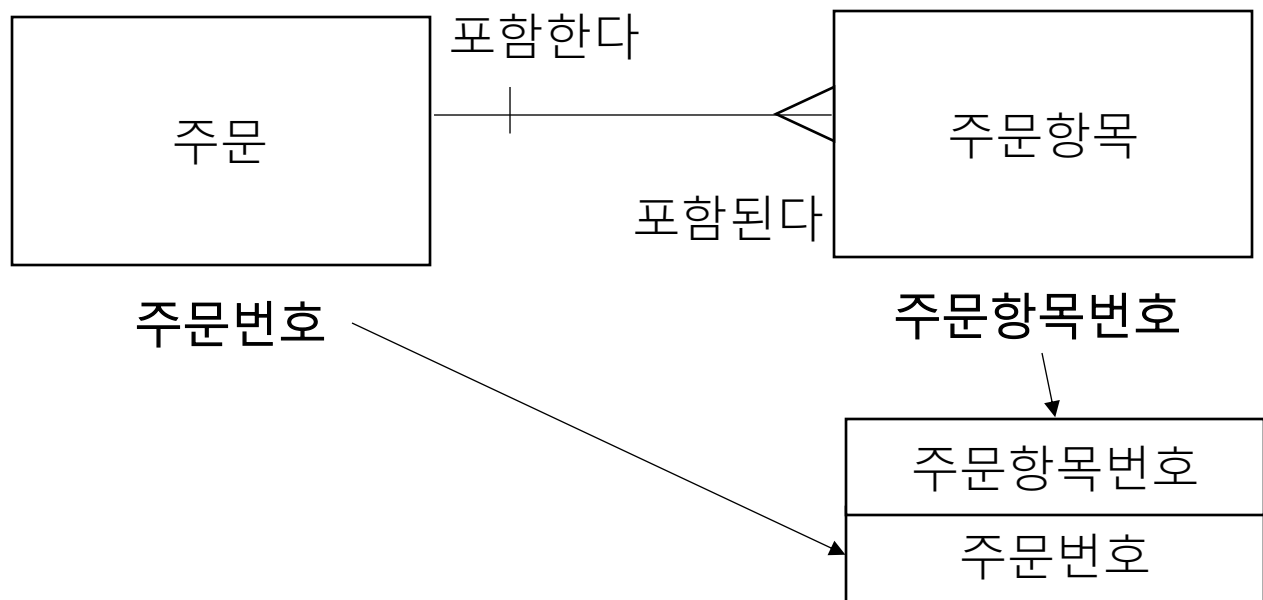
식별자의 구성-속성

사번	성명	주소	차량번호	결혼기념일
1111	이지아	서울시 양천구 신정동	041111	
2222	김효진	서울시 강남구 서초동	051111	

사번	성명	주소	차량번호	결혼기념일
1111	이지아	서울시 양천구 신정동	041111	
2222	김효진	서울시 강남구 서초동	051111	

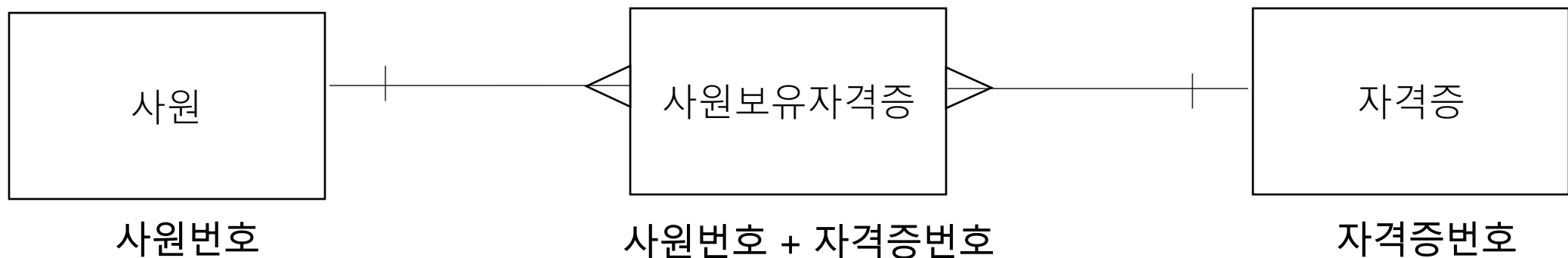
식별자의 구성 - 속성+관계

- 주문항목의 식별자
 - 주문항목번호(속성) + 포함된다[주문](관계)



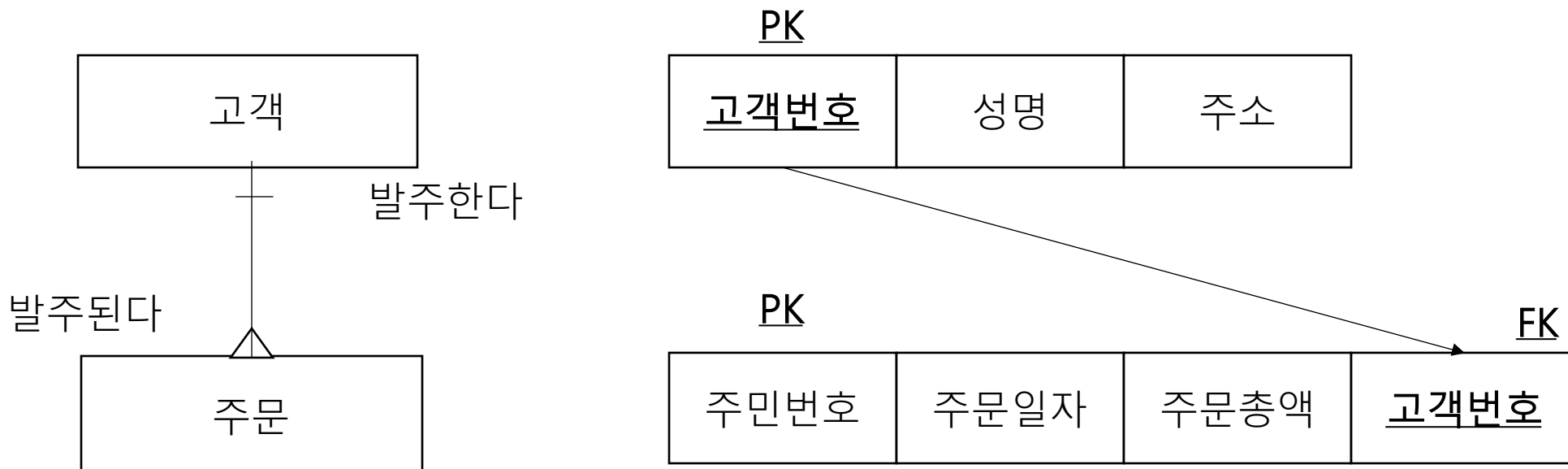
식별자의 구성 - 관계 + 관계

- 사원보유자격증의 식별자
 - 보유한다(사원) + 포함된다(자격증)



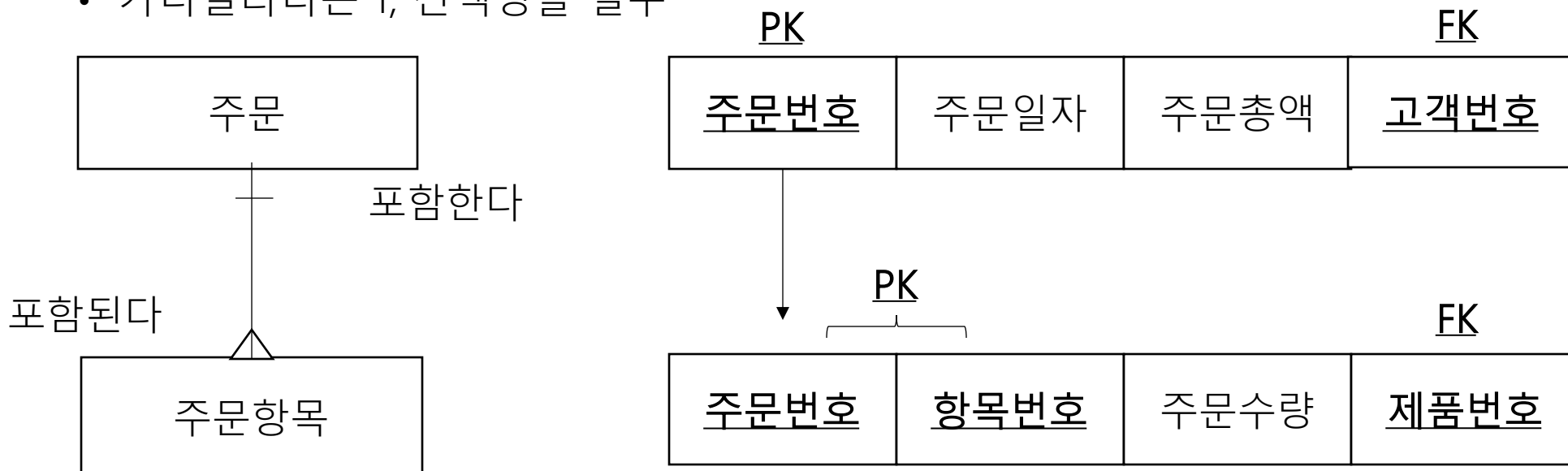
외부키(Foreign Key)

- 관계의 물리적 구현
 - 상위(Parent) 엔티티 타입의 식별자가 하위(Child) 엔티티 타입의 속성으로 구현



관계형 식별자

- 외부키가 식별자의 구성요소로 정의
 - 엔티티 타입의 계층적 구조 표현
 - 카디널리티는 1, 선택성을 필수



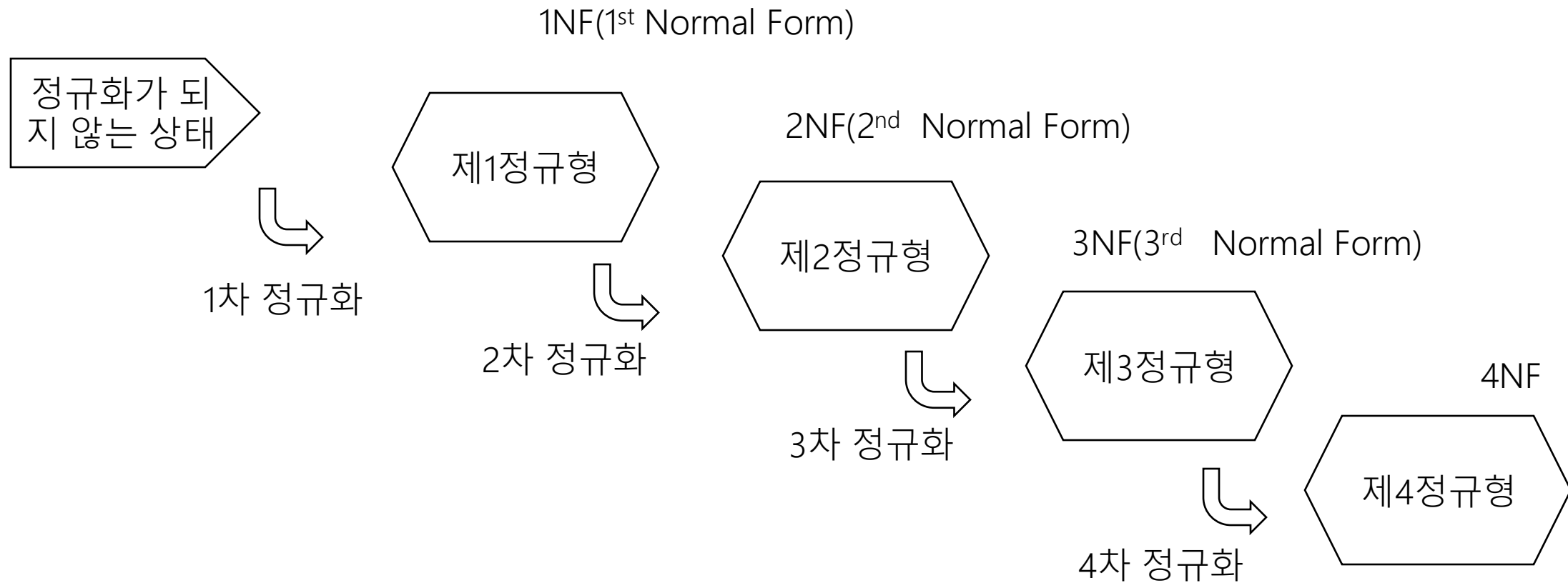
정규화 기법

- 효율적인 관리를 위한 데이터구조를 도출하기 위하여 엔티티 타입을 세분화하는 기법
 - 데이터의 종속성을 파악하여 독립적인 데이터는 서로 분리
- 정규화 효과
 - 데이터의 중복 최소화 - 데이터 저장공간의 최소화
 - 데이터 관리의 이상현상 발생 방지 - 정보 불일치, 손실 위험 최소화
 - 데이터 구조의 안정성 최대화

이상현상(Anomaly)

- 삽입 이상(Insert Anomaly)
 - 특정 엔티티를 삽입할 때 원하지 않는 속성값까지 삽입되는 현상
- 수정 이상(Update Anomaly)
 - 특정 엔티티의 속성값을 갱신 함으로써, 데이터의 일관성이 깨지는 현상
- 삭제 이상>Delete Anomaly)
 - 특정 엔티티를 삭제할 때까지 유지되어야 할 다른 속성까지 함께 삭제되는 현상

정규화 단계



최초의 상태

- 정규화가 수행되지 않는 상태
- 업무 규칙
 - 주문 접수 시에 고객번호별로 주문번호를 부여
 - 한번의 주문으로 여러 제품을 구매

주문

고객번호	주문번호	고객명	주문일자	주문총액	항목번호	제품번호
------	------	-----	------	------	------	------

제품명	제품단가	주문수량	주문금액
-----	------	------	------

제 1차 정규화

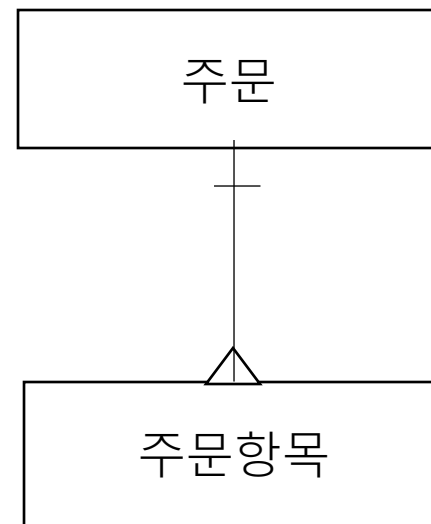
- 반복 그룹의 속성을 별도의 엔티티 타입으로 분리
 - 한 시점에 다수의 속성값을 가지는 속성

주문

고객번호	주문번호	고객명	주문일자	주문총액
------	------	-----	------	------

주문항목

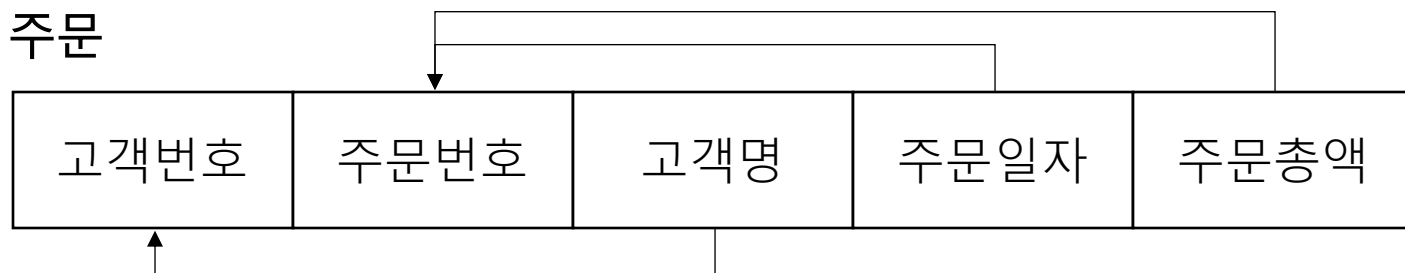
항목번호	제품번호	제품명	제품단가	주문수량	주문금액
------	------	-----	------	------	------



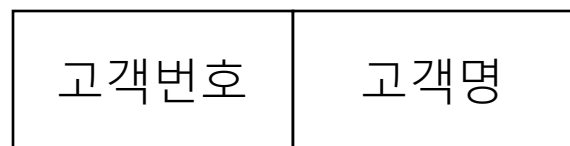
제2차 정규화

- 복합 식별자의 경우 전체에 종속적이지 않고 식별자 일부에만 종속적인 속성을 엔티티 타입으로 분리

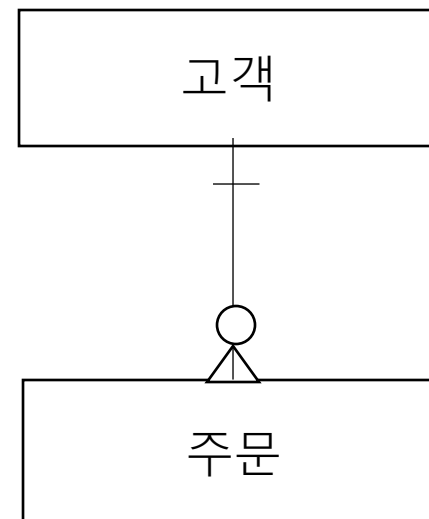
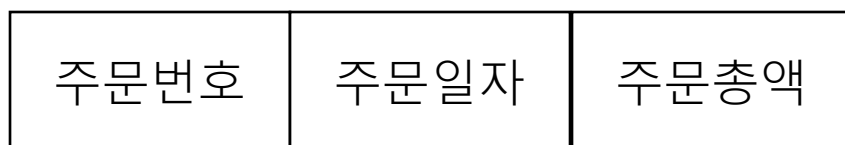
주문



고객



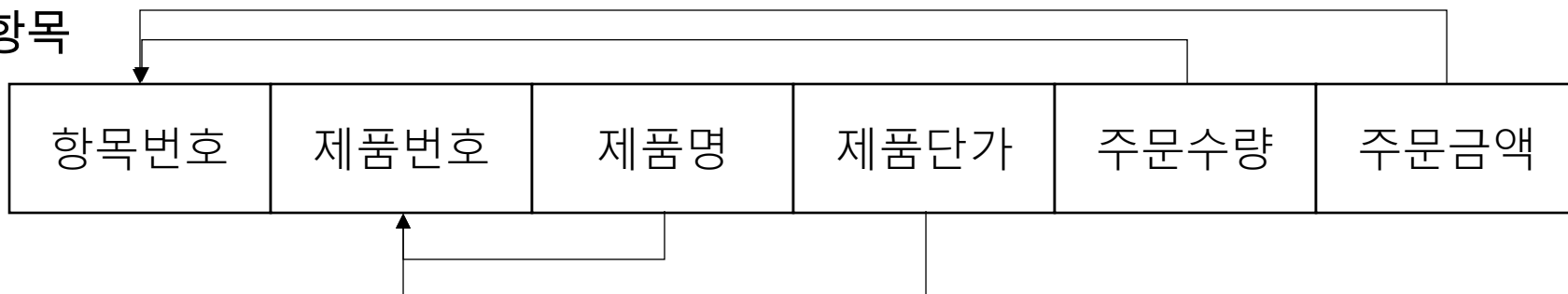
주문



제3정규화

- 식별자가 아닌 일반 속성에 종속적인 속성을 엔티티 타입으로 분리

주문항목

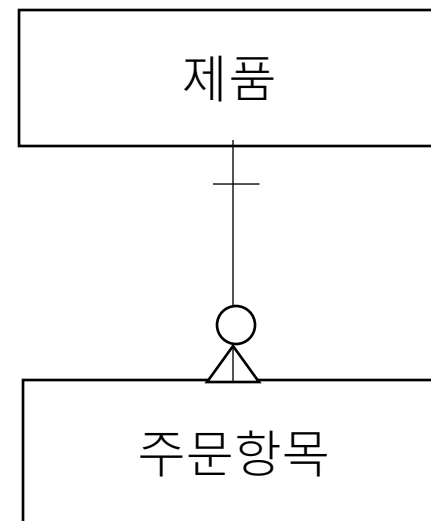


제품

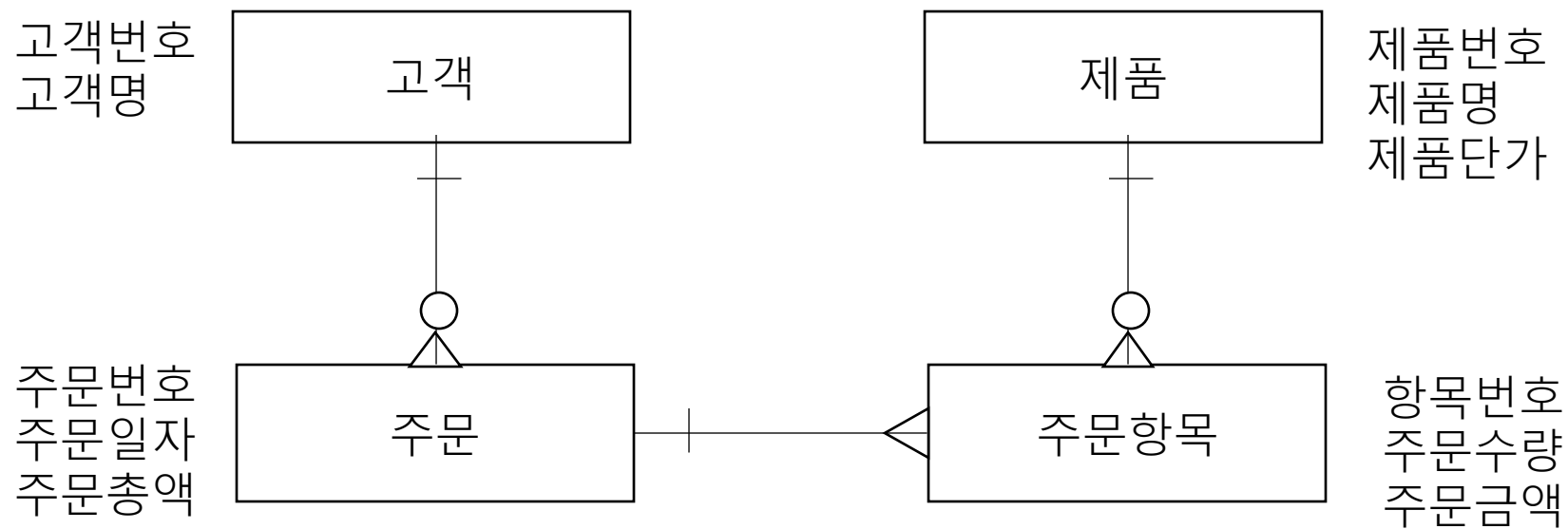
제품번호	제품명	제품단가
------	-----	------

주문항목

항목번호	주문수량	주문금액
------	------	------



정규화 결과



정규화 연습문제

- 3차 정규화까지 작업
- 업무규칙
 - 한 학생은 한 학기에 여러 과목 수강
 - 강사는 여러 과목을 강의
 - 강사 룸은 강사를 1인만 배정

학생

학번	학생명	학생주소	전공	과목코드	과목명	강사코드
----	-----	------	----	------	-----	------

강사명	강사룸	학점
-----	-----	----