



인 터 넷 진 화 의 열 쇠

온톨로지

웹 2.0에서 3.0으로

노상규·박진수 공저

god's toy business, 2007

SNU IDB laboratory

Ontology Contents

MODULE 1 온톨로지의 개념 및 응용

■ Chapter 1 온톨로지 개요

- 1. 온톨로지의 유래
- 2. 분류와 개념화 과정
- 3. 컴퓨터 온톨로지

■ Chapter 2 온톨로지의 분류와 용도

- 1. 온톨로지의 분류
- 2. 온톨로지의 사용 목적과 중요성
- 3. 온톨로지와 시맨틱 웹

■ Chapter 3 온톨로지 구축 프로젝트

- 1. 사이크(Cyc)
- 2. 워드넷(WordNet)
- 3. 전자거래문서
- 4. 통합의학언어시스템
- 5. 오픈 디렉터리 프로젝트
- 6. 국제상품분류코드(UNSPSC)

■ Chapter 4 온톨로지 적용 분야

- 1. 전자상거래 분야
- 2. 의료 분야
- 3. 법률 분야
- 4. 검색 서비스 분야
- 5. 문화컨텐츠 분야

MODULE 2 온톨로지 언어와 구축도구

■ Chapter 5 온톨로지 언어

- 1. 온톨로지 언어의 발전 과정
- 2. 인공지능 기반의 온톨로지 언어
- 3. 온톨로지 마크업 언어

■ Chapter 6 RDF(S): RDF와 RDF Schema

- 1. XML과 RDF
- 2. RDF
- 3. RDF Schema
- 4. RDF(S)의 한계점

■ Chapter 7 OWL(Web Ontology Language)

- 1. OWL의 기본 요소: 클래스와 속성
- 2. OWL의 새로운 기능
- 3. 세 종류의 OWL
- 4. OWL 예제

■ Chapter 8 토픽맵(Topic Maps)과 XTM(XML Topic Maps)

- 1. 토픽맵(Topic Maps) 개념
- 2. 토픽맵 구성요소
- 3. XTM 예제

■ Chapter 9 온톨로지 틀

- 1. 온톨로지 틀의 분류
- 2. 온톨로지 개발 틀
- 3. 주요 온톨로지 틀 요약 정보

Chapter 1 온톨로지의 개요

- 1. 온톨로지의 유래
- 2. 분류와 개념화 과정
- 3. 컴퓨터 온톨로지
 - 3.1 일반시스템 이론과 온톨로지
 - 3.2 번지 온톨로지(Bunge Ontology)
 - 3.3 온톨로지 정의
 - 3.4 온톨로지의 구성 요소

1. 온톨로지의 유래

- **Artificial Neuron**: Warren Mcculloch and Walter Pitts, 1943
- **“Artificial Intelligence”**: 1956 Dartmouth College, Workshop
- **Academic Founders of AI**
 - **John McCarthy (1927—2011)**
 - ▶ Stanford University, LISP
 - **Marvin Minsky (1927– Present)**
 - ▶ MIT, Perceptrons
 - **Allen Newell (1927 – 1992)**
 - ▶ Carnegie Mellon University, SOAR
 - **Herbert Simon (1916 – 2001)**
 - ▶ Carnegie Mellon University, General Problem Solver
 - **Alan Turing (1912 – 1954)**
 - ▶ University of Manchester, Turing Machine, Turing Test

1. 온톨로지의 유래

■ Types of Intelligence (red → ontology related intelligence)

- 논리적으로 추론하는 능력(reasoning)
- 추상적으로 사고하는 능력(abstraction)
- 계획하는 능력(planning)
- 문제를 해결하는 능력(problem solving)
- 개념과 언어를 이해하는 능력(comprehending ideas and languages)
- 학습하는 능력(learning)
- 사물을 분류하는 능력(grouping/classification)

1. 온톨로지의 유래

■ 온톨로지란?

- 존재 형태, 본질의 범주를 표현하는 철학분야: 존재론, 존재학
- 사물(thing)의 기본적인 범주나 세상구성 요소들의 개념을 다루는 학문 [Park, 2005]

■ 온톨로지에 관한 철학적 기반

- 고대 그리스 철학자: 플라톤과 아리스토텔레스 (기원전 4세기)
 - ▶ 현존하는 것에 나타나는 공통된 특징에 대한 질문에 답할수 있는 체계
 - ▶ 현실 세계를 단순화하려는 시도

■ 분류학(taxonomy)

- 사물이나 개념을 보다 잘 식별하여 이해하기 쉽게 배열하고 정리
- 일련의 표준을 제공함으로써 의사소통을 간소화
- 분류학이 잘 적용된 예(분야)
 - ▶ 테오프라스토스의 식물 분류표 → 현대 식물 분류학의 기초 예) 겉씨 식물, 속씨 식물
 - ▶ 칼 폰 린네의 이명법(Binominal Nomenclature) → 생물 분류학의 기초 예) Homo sapiens

Chapter 1 온톨로지의 개요

- 1. 온톨로지의 유래
- 2. 분류와 개념화 과정
- 3. 컴퓨터 온톨로지
 - 3.1 일반시스템 이론과 온톨로지
 - 3.2 번지 온톨로지(Bunge Ontology)
 - 3.3 온톨로지 정의
 - 3.4 온톨로지의 구성 요소

2. 분류와 개념화 과정

■ 분류(classification = taxonomy)

- 정보를 체계화하고, 그로부터 추론하여 결론 도출
- 지각하는 대상물에 관한 지식을 손쉽게 체계화
 - ▶ 존재하는 것들의 고유한 특징 → 특정한 범주(category)로 구분 가능
 - ▶ 관련된 다른 모든 범주와 연관시킴으로써 보다 쉽게 이해 가능
 - ▶ 정보 처리의 단순화 → 인식하는 대상을 보다 잘 이해
- 분류의 예
 - ▶ 의자 category: 의자들의 공통된 특징 (다리 3~4개, 등받이, 받침, 앉을수 있음)
 - ▶ 처음보는 어떤 객체 X (다리 3개, 받침)
 - ▶ X에 앉을 수 있다고 추론 가능

■ 개념화 (conceptualization) 단계들 → 분류

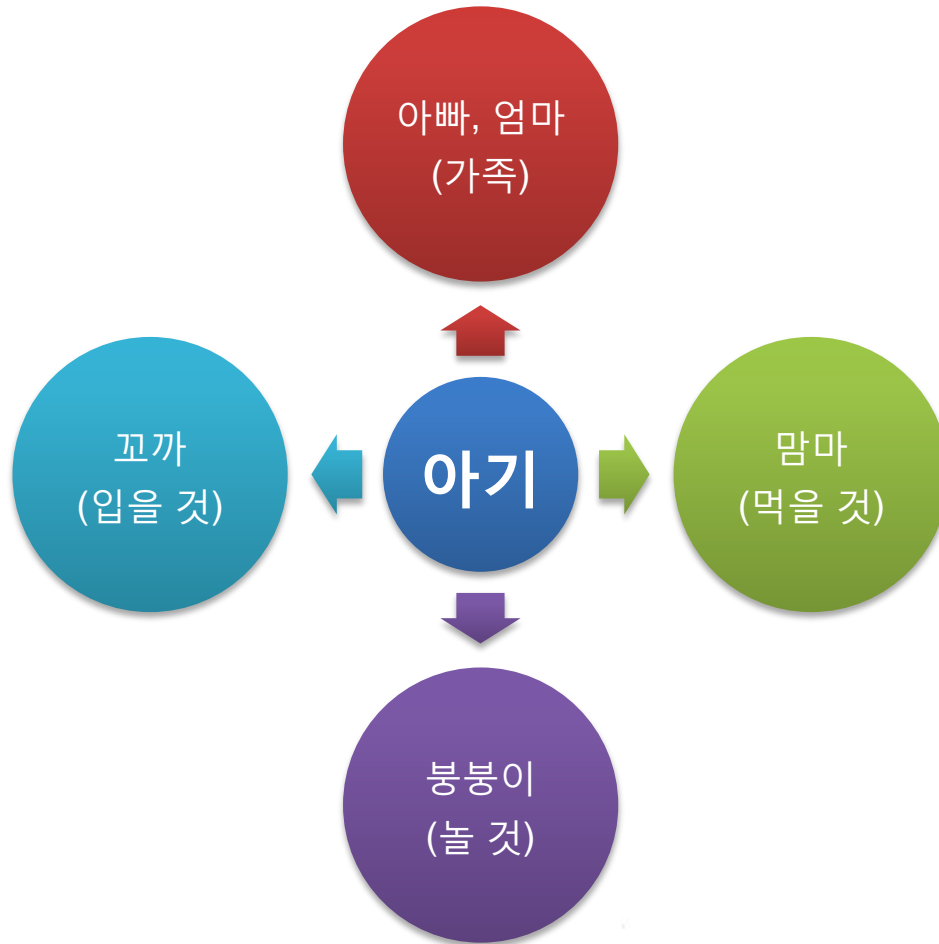
2. 분류와 개념화 과정

■ 개념화(conceptualization) 과정

- 인간은 태어나서 주변의 사물들을 인식하면서 분류
- 자신의 주위에 존재하는 것들을 무의식 중에 분류해 나가는 과정
- 개념화 과정의 예
 - ▶ 인간이 인식하는 세상에 대한 개념화 단계
 - 셔츠, 양말, 속옷 .. → 입을 것
 - 밥, 빵 .. → 먹을 것

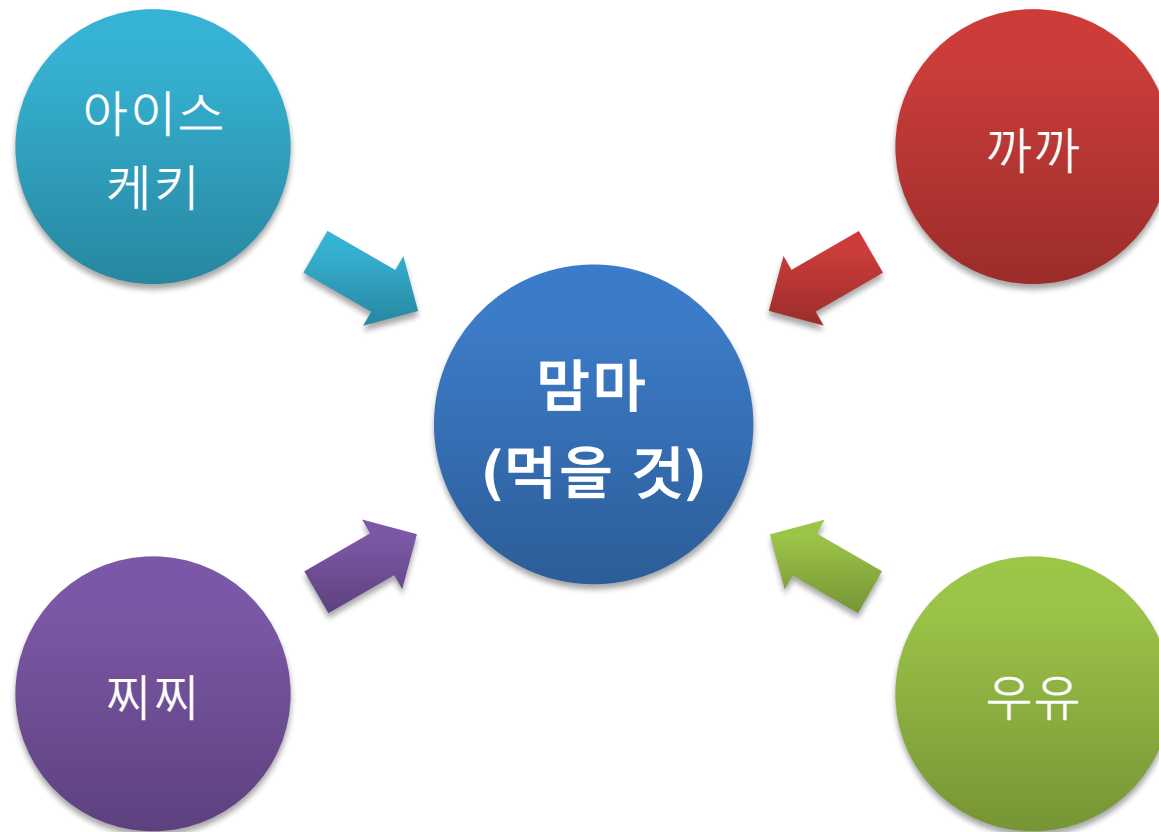
인간의 개념화 과정 [1/4]

- 아기가 인식하는 세계의 개념화 과정 1단계
 - 크게 4가지 정도의 category를 인지



인간의 개념화 과정 [2/4]

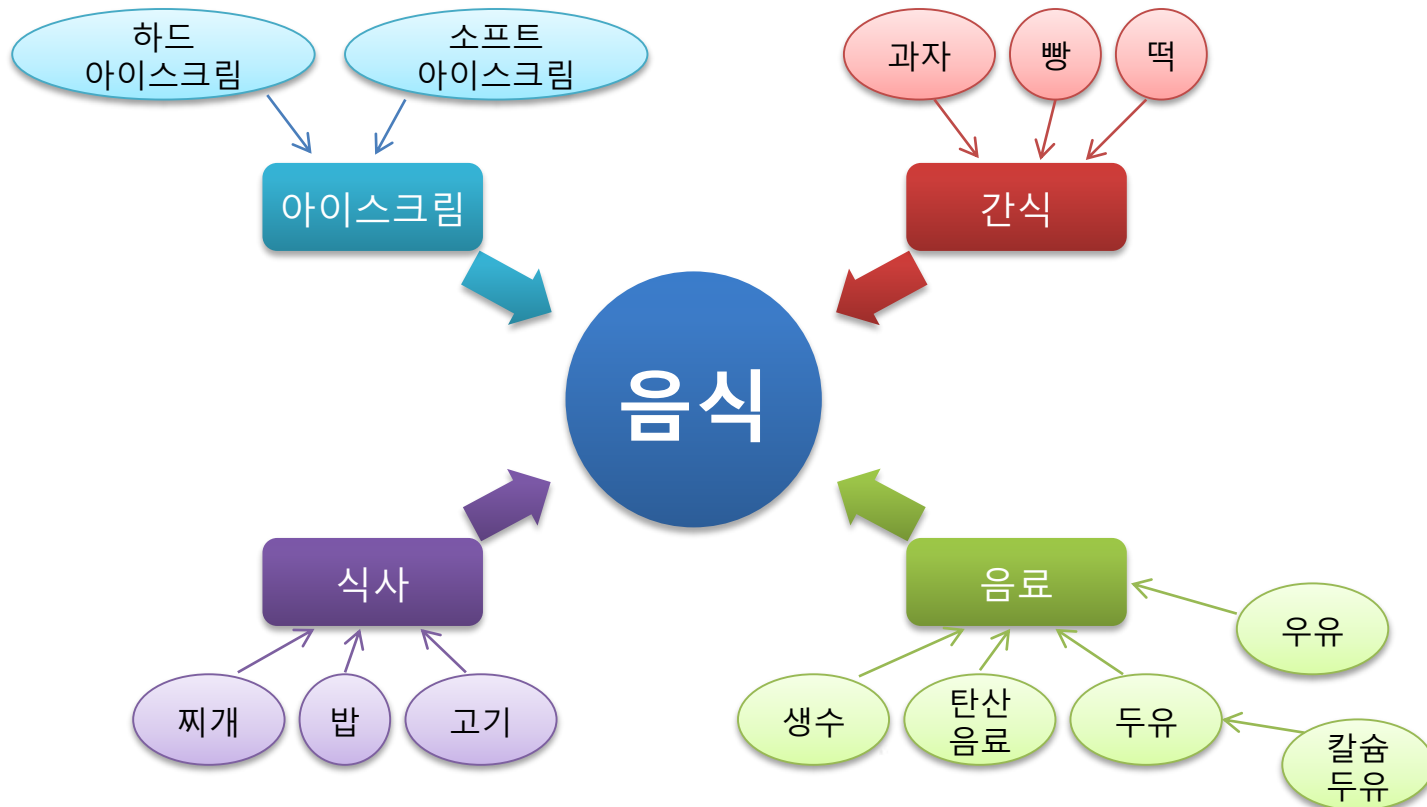
- 아기가 성장하며 인식하는 세계의 개념화 과정 2단계
 - 좀더 많은 것들을 접하면서 자신도 모르게 그것들을 분류



인간의 개념화 과정 [3/4]

■ 아기가 성장하며 인식하는 세계의 개념화 과정 3단계

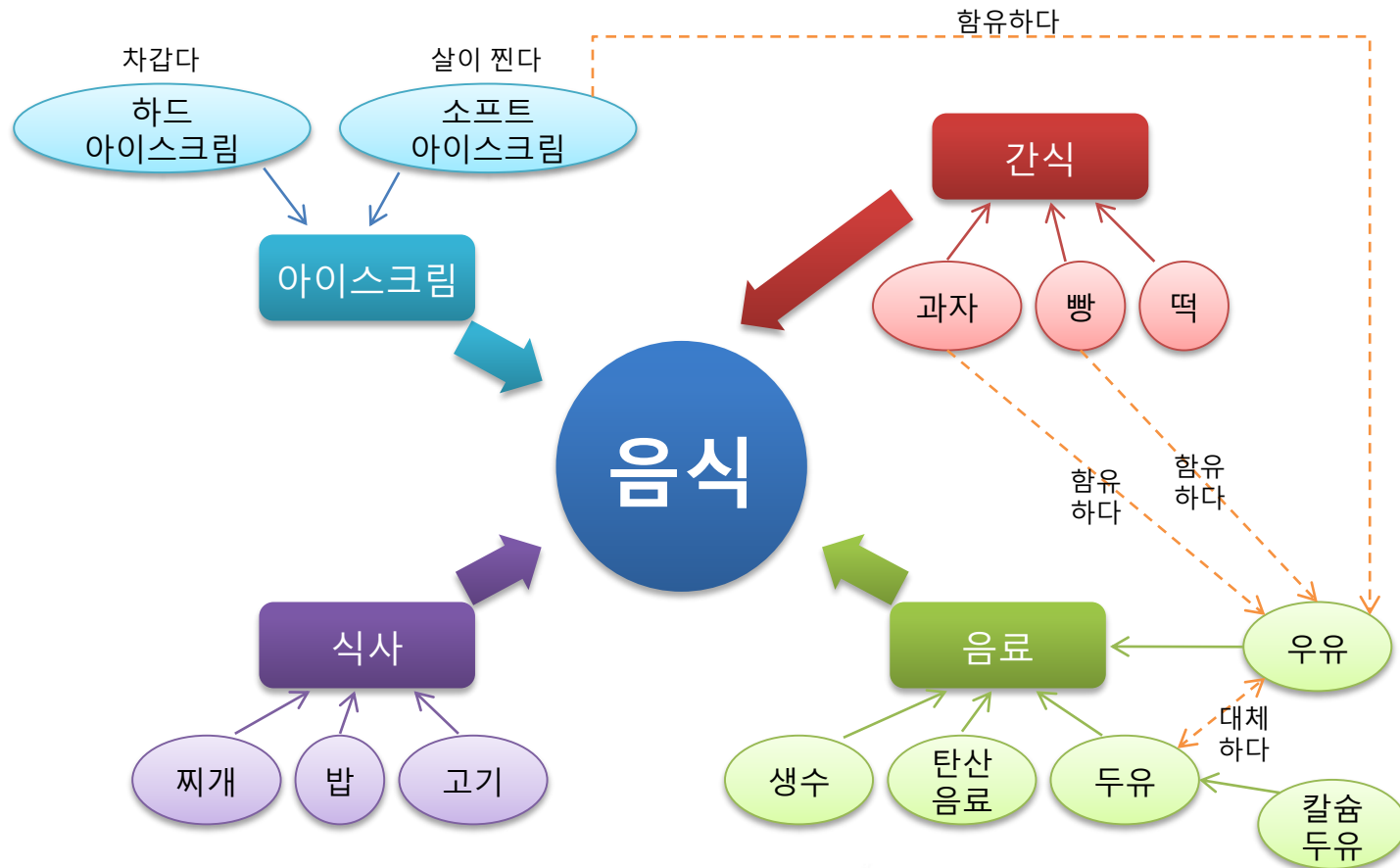
- 어린이가 되면서 먹을 것들의 종류가 더 많다는 것을 깨달음
- 먹을 것을 어떻게 분류하고 명명하는지 배움
 - ▶ 배우는 용어들은 사회 구성원들의 합의에 의해서 형성된 표준 용어



인간의 개념화 과정 [4/4]

■ 아기가 성장하며 인식하는 세계의 개념화 과정 4단계

- 각 범주에 속한 것들의 속성과 그들 간의 관계를 파악



2. 분류와 개념화 과정

■ 온톨로지

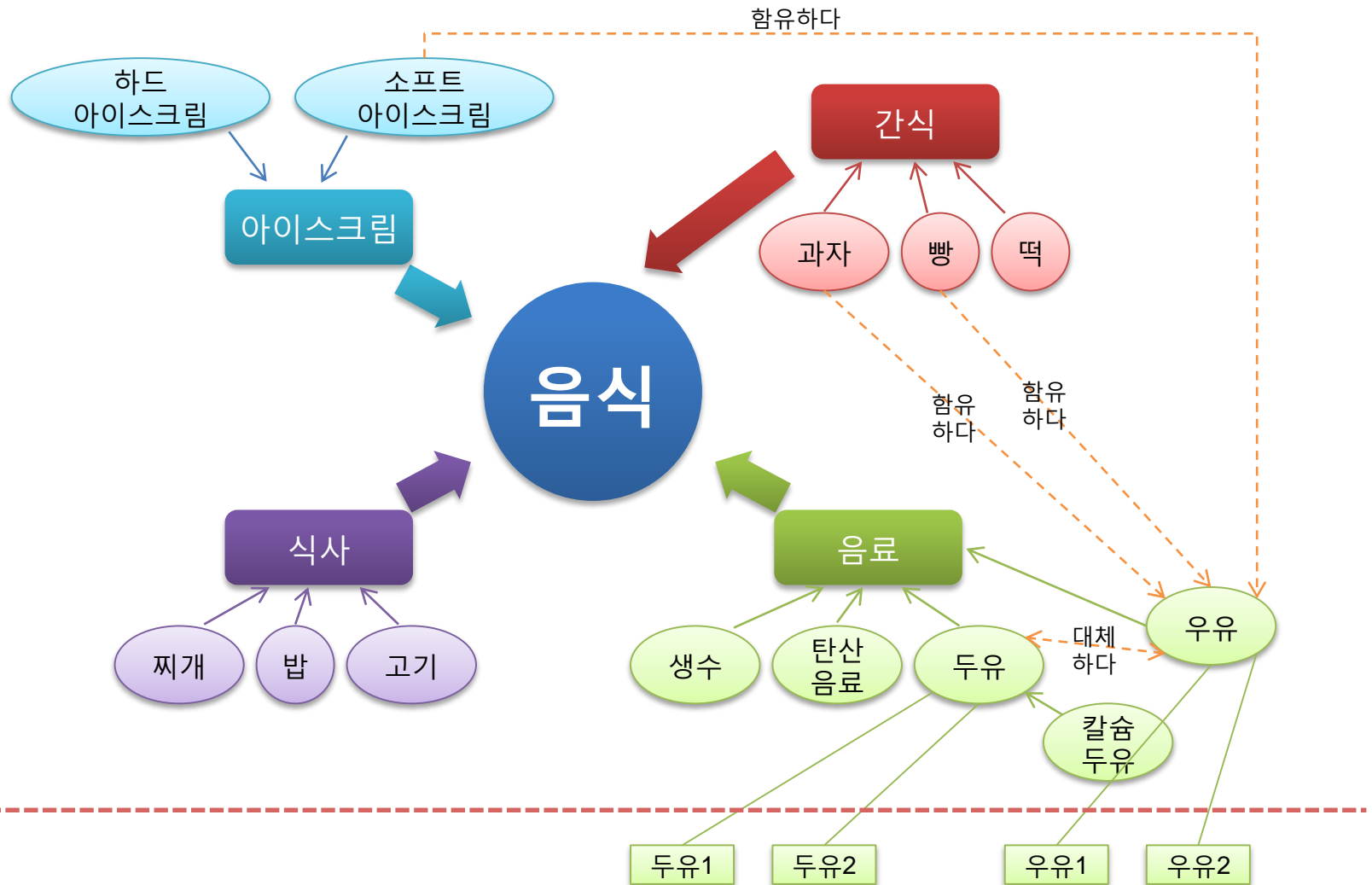
- 개념화 과정의 결과물
- 체계적 이론으로 세상의 모든 것들을 단순화, 추상화하여 분류
 - ▶ 계통을 세워 현상을 설명

■ 컴퓨터가 처리할 수 있는 형태로의 온톨로지 모델링

- 다양한 SW 분야에 적용 가능
- 예: 기업의 정보시스템에서의 온톨로지 응용
 - ▶ 비즈니스의 정보환경을 눈에 드러나도록 분명하게 모델링하여 체계화
 - ▶ 정보시스템들 간 정보 공유 및 처리를 자동화
 - ▶ 관련된 비즈니스와의 연동을 쉽게 함

2. 분류와 개념화 과정

■ 컴퓨터가 처리할 수 있는 형태의 온톨로지



Chapter 1 온톨로지의 개요

- 1. 온톨로지의 유래
- 2. 분류와 개념화 과정
- 3. 컴퓨터 온톨로지
 - 3.1 일반시스템 이론과 온톨로지
 - 3.2 번지 온톨로지(Bunge Ontology)
 - 3.3 온톨로지 정의
 - 3.4 온톨로지의 구성 요소

3. 컴퓨터 온톨로지

- 3.1 일반시스템 이론과 온톨로지
- 3.2 번지 온톨로지(Bunge Ontology)
- 3.3 온톨로지 정의
- 3.4 온톨로지의 구성 요소

3. 컴퓨터 온톨로지

- 특정 영역의 사물을 다룸 (Domain-Specific Ontology)
 - 특정 영역의 메타데이터를 컴퓨터가 해석할 수 있는 시맨틱으로 표현
- 기계가 이해하고 처리할 수 있도록 인간이 만든 인공물(Artifact)
 - 기계가 특정한 작업을 지능적으로 수행할 수 있게 해주는 핵심 기술
 - 인간과 이종(異種) 정보시스템 간의 일관성 있는 커뮤니케이션 수단

3.1 일반시스템 이론과 온톨로지

■ 시스템이란?

- 공통된 목적의 달성을 위해 여러 요소들이 상호작용을 하는 집합체
 - ▶ 예 : 기업은 이익 창출을 위해 자본, 노동력, 여러 프로세스로 이루어진 시스템
 - ▶ 넓게는 경제시스템, 경영시스템, 정치시스템, 교육시스템, 좁게는 기업, 학교, 가정 등
- 시스템의 대표속성: **창발적(emergent) 속성**
 - ▶ 이전에 없던 새로운 체계나 방법이 발생하는 특징
 - ▶ “전체는 부분의 합 이상이다.”
 - ▶ 시스템은 단순히 구성 요소의 합이 아니라 전체를 구성하는 요소들이 서로 연결되어 상호작용을 하는 집합체
 - ▶ 예) 기업의 매출액은 생산, 마케팅, 재무 등의 상호작용으로 나타나는 것이지 각 부서의 속성의 합이 아니다

3.1 일반시스템 이론과 온톨로지

■ 시스템의 기타 속성들

- 응집성: 시스템의 모든 구성 요소간의 일관성 있는 상호관련
- 패턴: 상호관련을 통해 시스템을 구조화
- 목적: 특정 목표를 달성하기 위해 목적이 있는 행동을 하는 속성
- 조절능력: 스스로 통제하고 조절
- 지속성: 살아남기 위해서 버티는 속성

■ 일반시스템 이론(general system theory) ➔ 현실세계의 Meta Model구축

- 일반적인 시스템과 관련된 타당한 원칙들을 도출하고 체계화
- 모든 시스템들에 보편적으로 적용될 수 있는 이론적 모델의 구축이 목적
- 시스템의 상호작용과 변화의 패턴을 설명하고 이해할 수 있는 틀 제공
- 시스템의 복잡성 이해를 용이
 - ▶ 상위시스템, 하위시스템의 계층 구조(hierarchical structure)로 표현

■ 어느 특정분야의 지식들의 시스템화 ➔ 온톨로지!!!!

3.2 Bunge 온톨로지

■ 마리오 번지 (Mario Bunge, 1919 ~)

- 아르헨티나 철학자
- Bunge 온톨로지 제안 (1977)

■ Bunge Ontology가 온톨로지를 이해하는데 도움이 되는 이유

- 온톨로지와 관련된 과거의 많은 연구에 기반하여 정립된 이론
- 현실 세계라는 시스템과 정보 시스템 모두에 적용 가능
- 수학적으로 잘 형식화된 개념, 전제, 일관성있는 표기
 - ▶ Thing, Attribute, Composite thing, Attribute, State, Event, Interaction, System

■ Ontology의 origin? → Maybe Ontology라는 단어를 처음 구사?

3.2 Bunge 온톨로지: 구성요소 [1/3]

■ **사물(thing):** 세상의 모든 구성 요소

- **속성(property):** 모든 사물은 고유한 속성을 지님
 - ▶ 하나의 사물에 본질적 or 여러 사물간에 상호적
 - ▶ 예) thing 돌멩이: property 무게, property 중력
- **합성물(composite thing)**
 - ▶ 사물들이 결합해 하나의 합성물을 형성할 수 있음
 - ▶ 컴퓨터 : 마이크로프로세서 + 메모리 + 하드디스크 등의 합성물
 - ▶ **창발적 속성(emergent property)**
 - 여러 사물들을 합성할 때 나타나는 속성
 - 프로세싱 파워 : 컴포넌트가 아닌 컴퓨터 시스템 전체가 지닌 속성

■ **애트리뷰트(attribute):** 인간이 사물에 부여한 특징

- ▶ 개개의 속성은 애트리뷰트로 나타낼 수 있음
- ▶ 예) '나이' : 인간의 성장과 노화 과정의 속성을 나타내는 애트리뷰트
- ▶ 사물의 속성 중 필요한 속성만을 애트리뷰트로 나타냄

3.2 Bunge 온톨로지: 구성요소 [2/3]

- **상태(state):** 어떤 특정 시점에 나타난 모든 애트리뷰트가 지닌 값의 집합
 - 자동차의 예
 - ▶ 애트리뷰트 : 총 주행거리, 연료량, 주행속도
 - ▶ 현실 세계의 변화는 존재하는 모든 것들의 상태 변화를 반영한 것
- **이벤트(event):** 상태의 변화를 일컫는 말
 - <초기 상태, 최종 상태, 변환>의 트리플(triple) 형태로 표현
 - 변환 : 변화에 영향을 주는 메커니즘

3.2 Bunge 온톨로지: 구성요소 [3/3]

- 상호작용(interaction): 시스템을 정의하는 데 있어서 중요한 개념
 - 만일 X와 Y중 적어도 어느 하나가 상대방에게 영향을 미친다면 X와 Y는 상호작용을 한다고 말함
- 시스템(system): 상호작용을 하는 것들로 구성
 - 합성(composition): 시스템 요소들의 집합
 - ▶ 합성물(composite thing) → 창발적 (emergent) 속성
 - 환경(environment): 시스템합성에 속하지 않는 요소들의 집합
 - 구조(structure): 시스템 요소들과 환경 요소들 간의 모든 상호작용의 집합
 - 하나의 사물 또는 사물들의 집합체(aggregate)
- 사물과 시스템의 역학(dynamics)
 - 상호작용(interaction)과 내적 변화(internal transformation)의 관점에서 설명 가능

3.3 온톨로지의 정의 [1/4]

- 20개도 넘는 컴퓨터 온톨로지(Computational Ontology)의 정의
 - [Bunge 1977], [Chandrasekran et al 1999], [Devedzic 2002], [Fensel 2001], 등등
 - “공유하는 개념화의 형식적이고 명확한 명세 [Thomas Gruber, 1993]”
 - ▶ 공유: 온톨로지가 합의된 지식을 표현해야 함
 - ▶ 개념화: 특정 영역의 현실 세계와 관련된 개념을 나타내는 추상 모델
 - ▶ 형식적: 컴퓨터가 읽고 처리 가능한 형태로 표현
 - ▶ 명확한: 특정 영역을 모델링에 필요한 개념들과 제약 조건들을 명시적으로 정의
- 컴퓨터가 해석·이해·처리할 수 있는 특정 영역의 지식체계
- 세상의 특정분야에 관련된 용어(개념)들을 정의하고, 이들간의 관계를 정리한 Dictionary

3.3 온톨로지의 정의 [2/4]

■ 개념 (concept)의 정의

- 어떤 사물에 대한 일반적이고 본질적 논리적 표현 (특정 언어에 비종속)
- 의미의 삼각 관계(meaning triangle): 사물과 개념, 기호간의 상호작용
 - 커뮤니케이션을 위해 대부분의 개념에는 기호가 부여(필수는 아님!)
- 개념과 기호의 matching issue
 - 같은 개념에 여러 기호가 부여 : ‘국가’ 와 ‘나라’
 - 다른 개념에 같은 기호가 부여 : 배 (운송 수단 / 과일 / 신체 일부 등)
 - 잘못된 커뮤니케이션 유발 가능성
- 합의된 온톨로지 → 바른 의미 전달 (커뮤니케이션)



3.3 온톨로지의 정의 [3/4]

■ 개념의 속성

- 내포(內包, intension)
 - ▶ 의미(meaning)
 - ▶ 어떤 개념의 내용이 되는 여러 속성
- 외연(外延, extension)
 - ▶ 어떤 개념이 적용되는 명제나 사물의 범위
- 예) 금강산
 - ▶ 금강산, 봉래산, 풍악산, 개골산
 - 서로 다른 내포(meaning) 을 가짐
 - ▶ 외연 : 각 명칭들은 모두 '금강산'이라는 외연을 지칭

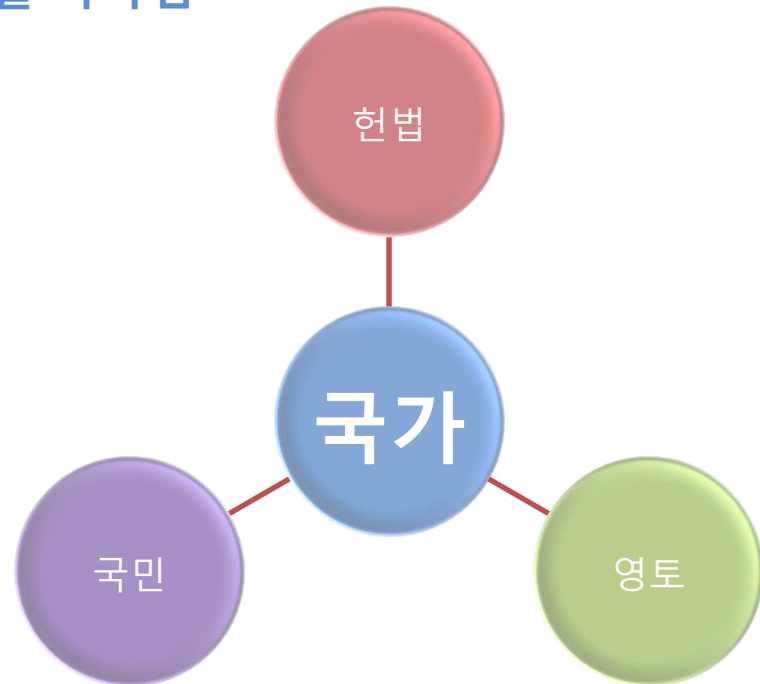
■ 온톨로지는 주로 내포와 관련

- ▶ 특정 용어와 상관없이 그 용어가 의도하는 개념 그 자체를 표현

3.3 온톨로지의 정의 [4/4]

■ 개념의 관계

- 개념은 다른 개념들과 연관
- 혼자 독자적으로 존재하는 경우는 거의 없음
 - ▶ '국가'는 '국민', '헌법', '영토' 등의 개념과 상관관계
- 온톨로지는 개념과 개념간의 관계를 나타냄
→ 개념의 의미를 명확히 함



3.4 온톨로지의 기본 요소

- 개념(concept)
- 속성(property)
- 관계(relationship)
- 제약조건(constraint)
- 공리(公理, axiom)
- 인스턴스(instance)

3.4 온톨로지의 구성 요소 [1/5]

■ 개념(concept)

- 현실 세계에서 사물(thing)에 대한 일반적이고 본질적인 인식이나 지식
- 물리적, 추상적인 것 모두 포함: '사랑', '우정', '천사' 등
- 단순개념(elementary concept): 더 이상 나누어지지 않는 기본 개념
- 합성개념(composite concept)
 - ▶ 단순개념들이 연관성을 가지고 결합되어 이루어진 개념
 - ▶ 예) 컴퓨터 = CPU + 메모리 + 하드디스크 등 다수의 개념으로 성립

■ 속성(property)

- 개념에 근본적으로 속해 있는 성질
- '프린터'라는 개념은 '가격', '색깔', '무게', '제조회사' 등과 같은 속성을 가짐
- 애트리뷰트(attribute)와 유사

3.4 온톨로지의 구성 요소 [2/5]

■ 관계(relationship)

- 개념들 사이의 상관관계
- 상속관계(inheritance relationship: ISA)
 - ▶ 온톨로지 개념 간의 계층구조(hierarchical structure) 형성
 - 예) '레이저 프린터' ISA '프린터'
- 합성관계(composite relationship: IS-PART-OF)
 - ▶ 상속관계 이외의 계층구조 형성
 - ▶ 부분전체관계(part-whole relationship)
 - 부분(is-part-of)관계로 표현: 'CPU', '메모리', '하드디스크'는 '컴퓨터'와 부분관계로 표현

■ 제약조건(constraint)

- 개념들 간의 관계나 속성의 값에 관한 제한 규정
 - ▶ 컴퓨터와 CPU의 관계: 컴퓨터는 반드시 1개 이상의 CPU를 가짐
 - ▶ 프린터의 속성인 가격: 반드시 0 이상의 정수값을 가짐
- 지식을 표현하거나 추론을 할 때 유용

3.4 온톨로지의 구성 요소 [3/5]

■ 공리(公理, axiom)

- 추론의 기본이 되는 명제
- 증명을 할 수 없거나 증명을 요하지 않는 ‘참(true)’으로 인정되는 문장
 - ▶ “프린터는 복사용지를 사용한다.”
 - ▶ “삼성Lightning은 삼성LPT100 레이저 토너를 사용한다.”
 - ▶ “삼성Lightning과 HPSpeed0는 상호 대체한다.”
- 모든 공리는 제약조건
- 논리적 정확성을 검증하거나 새로운 사실을 추론할 때 유용하게 사용

■ 인스턴스(instance)

- ‘개체’라고도 하며, 각 개념의 실례를 의미
 - ▶ ‘프린터1’, ‘프린터2’는 ‘삼성Lightning’의 인스턴스
 - ▶ ‘토너1’, ‘토너2’, ‘토너3’은 ‘삼성LPT100’의 인스턴스

■ 컴퓨터 관련 기기 및 사무용품 온톨로지



3.4 온톨로지의 구성 요소 [5/5]

■ 온톨로지의 활용 → Smart Operation, Intelligent Operation

● 인터넷 쇼핑몰 검색 시스템

- ▶ 여러 회사의 제품을 비교하고 상품들의 관계를 정리
- ▶ 고객: 자신이 알고 있는 정보, 사고자 하는 제품과 직접적/간접적인 정보를 제공
 - 원하는 제품을 손쉽게 찾을 수 있음
- ▶ 사례 : 고객 '태희'의 구매 의도
 - 사고자 하는 것은 컬러 레이저 프린터이다.
 - 기존에 쓰던 제품은 HP에서 나온 Speed0라는 제품이다.
 - 사고자 하는 제품의 브랜드는 삼성이다.
 - 기존의 제품과 기능이나 가격 면에서 비슷한 제품을 구입하고자 한다.