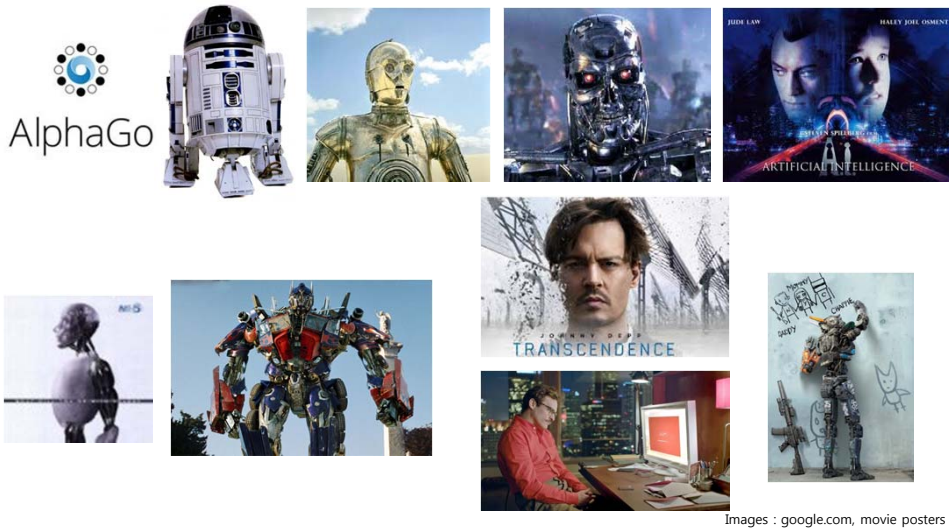


인공지능의 소개

충북대학교 소프트웨어학과
이건명

- ❖ 인공지능
- ❖ 인공지능의 역사
- ❖ 요소 기술 분야
- ❖ 주요 응용 분야
- ❖ 최근 동향
- ❖ 인공지능의 윤리
- ❖ 인공지능의 특이점

1. 인공지능



1. 인공지능

- ❖ 지능(知能, intelligence)
 - 본능적이나 자동적으로 행동하는 대신에, **생각하고 이해하여 행동하는 능력**
- ❖ 인공지능(人工知能, Artificial Intelligence)
 - 인공적으로 만든 지능
 - **튜링 테스트**(Turing test)
 - 지능의 조작적 정의 (operational definition)
 - 조작적 정의: 측정할 수 있는 조건으로 어떤 속성을 기술



Alan Mathison Turing
(1912.6 - 1954.6)

튜링 상(Turing Award)

ACM에서는 매년 컴퓨터 분야의 기여가 큰 연구자에게 시상.
\$250,000 (Intel, Google 후원)

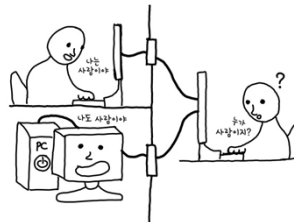


Image : wikipedia

인공지능

❖ 인공지능(Artificial Intelligence)

- **Dartmouth Conference** (1956, Dartmouth University, USA)
 - John McCarthy가 AI 용어 제안



(1927-2011)



- **AI@50 conference** (2005) : Trenchard More, John McCarthy, Marvin Minsky, Oliver Selfridge, Ray Solomonoff

Image : www.dartmouth.edu, wikipedia

인공지능

❖ 인공지능(人工知能, Artificial Intelligence)

- 사람의 생각과 관련된 활동, 예를 들면 의사 결정, 문제 해결, 학습 등의 활동을 자동화하는 것 (Bellman, 1978)
- 사람이 하면 더 잘 할 수 있는 일을 컴퓨터가 하도록 하는 방법을 찾는 학문 (Rich & Knight, 1991)
- 지능이 요구되는 일을 할 수 있는 기계를 만드는 예술 (Kurzweil, 1990)
- 지능적인 에이전트를 설계하는 학문 (Poole et al. 1998)
- 인지하고, 추론하고, 행동할 수 있도록 하는 컴퓨팅에 관련된 학문 (Wilson, 1992)
- 인공물이 지능적인 행위를 하도록 하는 것 (Nissson, 1990)

강한 인공지능과 약한 인공지능

❖ 강한 인공지능 (strong AI)

- 사람과 같은 지능
- 마음을 가지고 사람처럼 느끼면서 지능적으로 행동하는 기계
- 추론, 문제해결, 판단, 계획, 의사소통, 자아 의식(self-awareness), 감정(sentiment), 지혜(sapience), 양심(conscience)
- 튜링 테스트

❖ 약한 인공지능 (weak AI, narrow AI)

- 특정 문제를 해결하는 지능적 행동
- 사람의 지능적 행동을 흉내 낼 수 있는 수준
- 대부분의 인공지능 접근 방향
- 중국인 방 사고실험(Chinese room thought experiment)

강한 인공지능과 약한 인공지능

❖ 중국인 방 사고실험(The Chinese Room Thought Experiment)



(1932-)

- John Searle (1980) 제시
 - 문 밑으로 중국어로 쓴 질문지를 전달
 - 방 안에서 중국어를 모르는 사람이 글자모양에 따른 중국어 단어 조합 방법 매뉴얼을 참조하여 답변에 대한 단어 조합
 - 조합된 단어들을 문 밖으로 내보냄
 - 문 밖 사람은 중국어를 이해하는 사람이 방안에 있다고 생각
 - 단지 흉내만 내고 이해하는 것은 아님
- 이해하지 못하고 흉내 낼 수 있어도 지능적(intelligent) 행동



If you see this shape,	then produce this shape,
"什麼"	"爲天"
followed by this shape,	followed by this shape,
"帶來"	"下式".
followed by this shape,	
"快樂"	

Image : wikipedia, http://www.mind.ilstu.edu/curriculum/searle_chinese_room/searle_chinese_room.php



2. 인공지능의 역사

❖ 1960년대 이전

- 1946년 펜실베니아 대학, ENIAC 개발
- 큰 기대와 여러 가지 시도, 매우 제한된 성공
- LISP 언어 개발 (McCarthy, 1958)
- **Advice Taker** – 공리(axiom) 기반 지식표현 및 추론 (McCarthy)
- **Perceptron** – 신경망 모델 (Rosenblatt, 1958)
- 수단-목표 분석(means-ends analysis) 기법 (Newell & Simon, 1958)
 - 범용 문제해결을 목표로 한 GPS(General Problem Solver) 개발



Frank Rosenblatt
(1928-1971)



Allen Newell
(1927-1992)
전산학, 심리학



Herbert Simon
(1927-1992)
경제학 노벨상, 1978

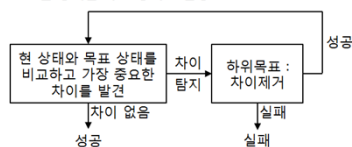
Image : wikipedia

인공지능의 역사

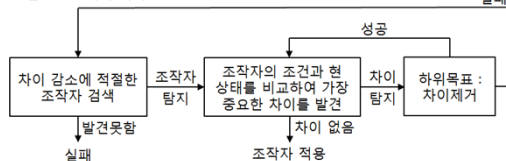
❖ 수단-목표 분석(means-ends analysis)

- 해결해야 하는 문제를 **상태(state)**로 정의
- **현재 상태**와 **목적 상태(goal state)** 간의 **차이** 계산
- 목적 상태로 도달하기 위한 **조작자(operator, 연산자)** 선택 적용하는 과정 반복

흐름도 I : 현 상태를 목표 상태로 변경



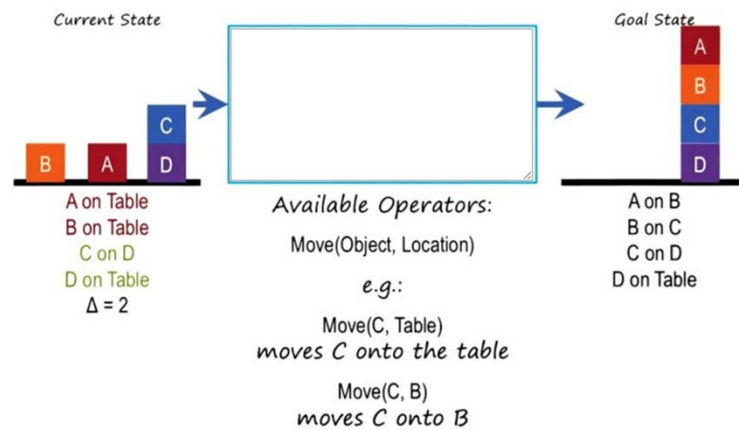
흐름도 II : 차이 제거



인공지능의 역사

❖ 수단-목표 분석(means-ends analysis) – cont.

- 예. 블록이동 문제



Source: udacity

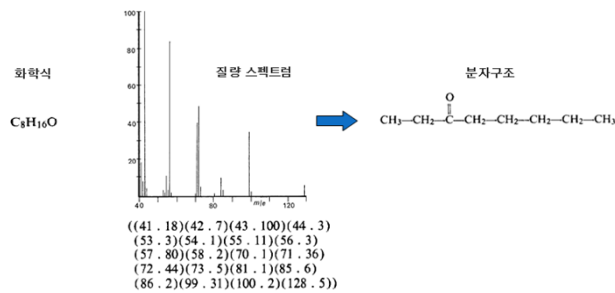
인공지능의 역사

❖ 1970년대에서 1980년대 초반

- 일반적인 방법보다는 **특정 문제 영역**에 효과적인 방법을 찾는 연구
- **전문가 시스템**(expert system)
 - 특정 영역의 문제에 대해서는 전문가 수준의 해답을 제공
 - 1970년대 초반부터 1980년대 중반 상업적 성공 사례
 - **MYCIN, PROSPECTOR, DENDRAL**
- **전문가 시스템 개발 도구**(expert system shell) 개발
- **Prolog** 개발
 - 지식의 표현과 추론을 지원하는 **논리**(logic) 기반 언어

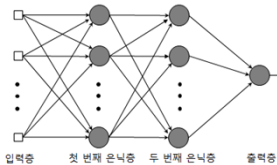
인공지능의 역사

- 대표적인 초창기 전문가 시스템
 - MYCIN
 - 스탠포드 대학에서 개발한 전염성 혈액 질환 진단
 - PROSPECTOR
 - 광물탐사 데이터 분석
 - DENDRAL
 - 화학식과 질량 스펙트럼 데이터로부터 유기화합물의 분자구조 결정
 - 스탠포드 대학의 Edward Feigenbaum 팀 개발



인공지능의 역사

- ❖ 1980년대 중반에서 1990년대
 - 신경망 모델 발전
 - 다층 퍼셉트론(multi-layer Perceptron, MLP) : 신경망의 르네상스
 - 오차 역전파(error backpropagation) 알고리즘



- 퍼지이론(fuzzy theory)
- 진화연산 : 유전자 알고리즘, 진화 프로그래밍
- 확률적 그래프 모델
 - 베이저안 네트워크(Bayesian network)
 - 마르코프 랜덤 필드(Markov random field)
- 서포트 벡터 머신(Support Vector Machine, SVM)

인공지능의 역사

❖ 2000년대 이후

- 에이전트(agent)
- 시맨틱 웹(semantic web)
- 기계학습(machine learning)
- 데이터 마이닝(data mining)
- 심층 학습(deep learning, 딥러닝)
- 상업적 성공 사례 다수 출현
 - 애플 Siri, 구글 Now, MS Cortana
 - IBM의 Watson
 - 로봇 : Boston Dynamics 군사용 로봇 BigDog, 재난구조 로봇 Atlas
 - 자율주행 자동차
 - 클라우드 앱 서비스

인공지능의 역사

- 튜링 테스트 최초 통과 프로그램
 - Eugene Goostman
 - 13세인 우크라이나 소년 캐릭터의 챗봇(chatbot)
 - 튜링 사망 60주년을 기념하여 진행된 이벤트에서 튜링 테스트 통과 (2014.6)
 - » 30명의 검사자가 5분씩 대화: 33%가 사람으로 판정
 - » 제한적 상황에 대한 통과



Scott: Which is bigger, a shoebox or Mount Everest?
 Eugene: I can't make a choice right now. I should think it out later. And I forgot to ask you where you are from...
 Scott: How many legs does a camel have?
 Eugene: Something between 2 and 4. Maybe, three? :-))) By the way, I still don't know your specialty - or, possibly, I've missed it?

인공지능의 역사

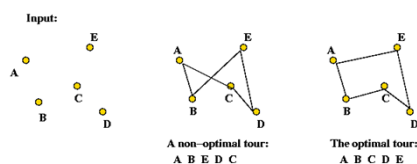
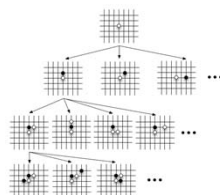
- 인공지능 바둑 프로그램 알파고(AlphaGo)
 - 구글 DeepMind 개발
 - 2016.3.9~3.15 총 5회의 대국에서 알파고가 4승 1패로 승리
 - 기계학습과 병렬처리로 구현



3. 연구 지능의 요소 기술

❖ 탐색(search)

- 문제의 답이 될 수 있는 것들의 집합을 **공간(space)**으로 간주하고, 문제에 대한 **최적의 해**를 찾기 위해 공간을 **체계적으로 찾아 보는 것**
- 무정보 탐색**
 - 너비우선 탐색(breadth-first search), 깊이우선 탐색(depth-first search)
- 휴리스틱 탐색**
 - A* 알고리즘
- 게임 트리 탐색**
 - mini-max 알고리즘, α - β 가지치기(pruning), 몬테카를로 트리 탐색



연구 지능의 요소 기술

❖ 지식표현(knowledge representation)

- 문제 해결에 이용하거나 심층적 추론을 할 수 있도록 지식을 효과적으로 표현하는 방법
- IF-THEN 규칙(rule)
- 프레임(frame)
- 의미망(semantic net)
- 논리(logic) : 명제논리(propositional logic), 술어논리(predicate logic)
- 스크립트
- 온톨로지 기술 언어 : RDF, OWL
- 불확실한 지식 표현 : 확신도, 확률기반 표현, 퍼지 이론
- 확률 그래프 모델
- 함수 기반 지식표현

연구 지능의 요소 기술

❖ 추론(inference)

- 가정이나 전제로부터 결론을 이끌어내는 것
- 관심 대상의 확률 또는 확률분포를 결정하는 것
- 규칙기반 시스템의 추론
 - 전향추론(forward inference)
 - 후향추론(backward inference)
- 확률 모델의 추론
 - 베이즈 정리(Bayesian theorem)
 - 주변화(marginalization)



$$P(A) = \sum_b P(A, B = b)$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

증거

사후확률 가능성도 사전확률

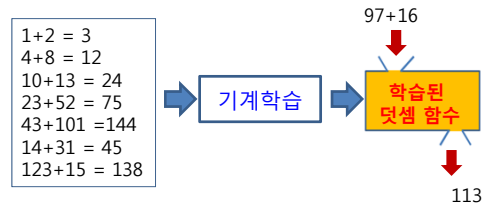
연구 지능의 요소 기술

❖ 기계 학습(machine learning)

- 경험을 통해서 나중에 유사하거나 같은 일(task)를 더 효율적으로 처리할 수 있도록 시스템의 구조나 파라미터를 바꾸는 것
- 알고 있는 것으로부터 모르는 것을 추론하기 위한 알고리즘을 만드는 것
- 지도학습, 비지도학습, 강화학습

▪ 지도학습(supervised learning)

- 문제(입력)와 답(출력)의 쌍으로 구성된 데이터들이 주어질 때, 새로운 문제를 풀 수 있는 함수 또는 패턴을 찾는 것

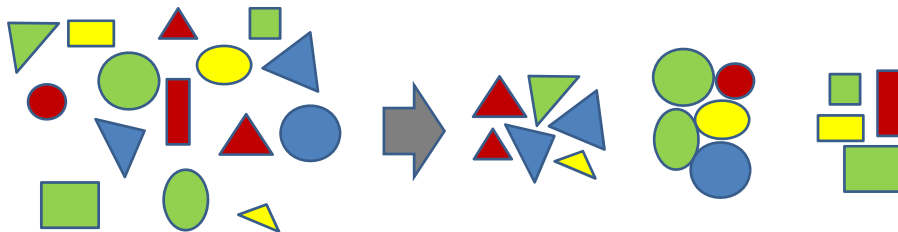


연구 지능의 요소 기술

❖ 기계 학습(machine learning) – cont.

▪ 비지도학습(unsupervised learning)

- 답이 없는 문제들만 있는 데이터들로부터 패턴을 추출하는 것



연구 지능의 요소 기술

❖ 기계 학습(machine learning) – cont.

▪ 강화학습(reinforcement learning)

- 문제에 대한 직접적인 답을 주지는 않지만 경험을 통해 **기대 보상(expected reward)**이 최대가 되는 **정책(policy)**을 찾는 학습

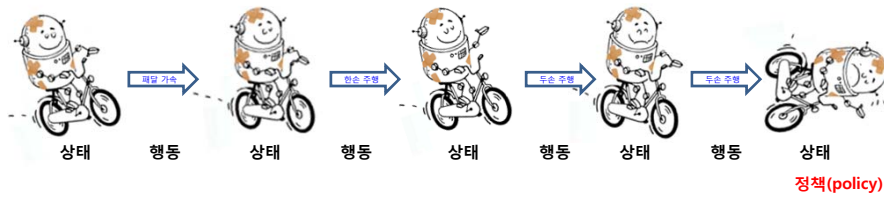


Image source : RÉMI MUNOS
RAPPORT ANNUEL INRIA 2008

연구 지능의 요소 기술

❖ 계획수립(planning)

- 현재 상태에서 **목표하는 상태**에 도달하기 위해 수행해야 할 일련의 **행동 순서**를 결정하는 것
- 작업 수행 절차 계획
- 로봇의 움직임 계획



<http://www.youtube.com/watch?v=1u4PR-pfghU>

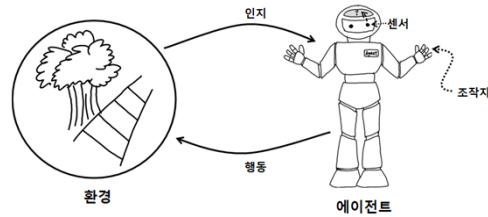


Image : spiral, gamma.cs.unc.edu

연구 지능의 요소 기술

❖ 에이전트(agent)

- 사용자로부터 위임받은 일을 자율적으로 수행하는 시스템
 - BDI(Belief-Desire-Intention) 모델
 - Belief (환경에 대한 정보), Desire(목적), Intention(의도, 목적 달성을 위한 세부 전술적 목표)

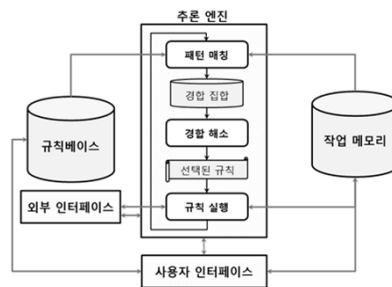


- 소프트웨어 에이전트(software agent)
 - 채팅봇(chatting bot) : Siri, Tay
- 물리적 에이전트(physical agent)
 - 로봇(robot)

4. 주요 응용 분야

❖ 지식기반 시스템(knowledge-based system)

- 지식을 축적하고, 이를 이용하여 서비스를 제공하는 시스템
- 전문가시스템(expert system)
 - 특정 문제 영역에 대해 전문가 수준의 해법을 제공하는 것
 - 간단한 제어시스템에서부터 복잡한 계산과 추론을 요구하는 의료 진단, 고장 진단, 추천 시스템 등

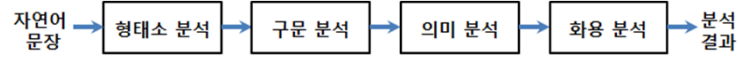


규칙기반 전문가시스템 구조

주요 응용 분야

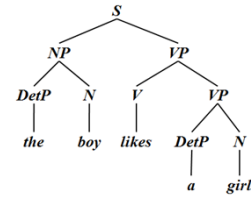
❖ 자연어 처리(natural language processing)

- 사람이 사용하는 일반 언어로 작성된 문서를 처리하고 이해하는 분야



친구에게서 : 친구(명사) + 에게(조사) + 서(조사)

[자베르]인형 경감이 [장발장]인형과 [1832]불타는 [파리]저항에서 마주쳤다



- 형태소 분석, 구문분석, 품사 태깅, 의미분석
- 언어모델, 주제어 추출, 객체명 인식
- 문서 요약, 기계번역
- 질의 응답

주요 응용 분야

❖ 데이터 마이닝(data mining)

- 실제 대규모 데이터에서 암묵적인, 이전에 알려지지 않은, 잠재적으로 유용할 것 같은 정보를 추출하는 체계적인 과정



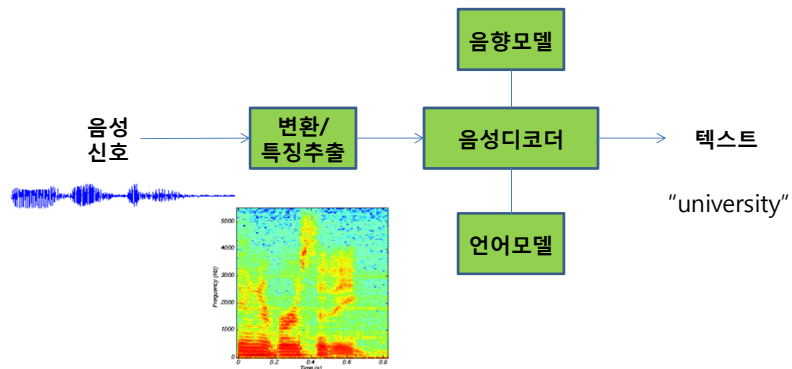
- 연관 규칙
- 분류 패턴, 군집화 패턴
- 텍스트 마이닝
- 그래프 마이닝
- 추천
- 시각화(visualization)

www.businessnewsdaily.com

주요 응용 분야

❖ 음성 인식

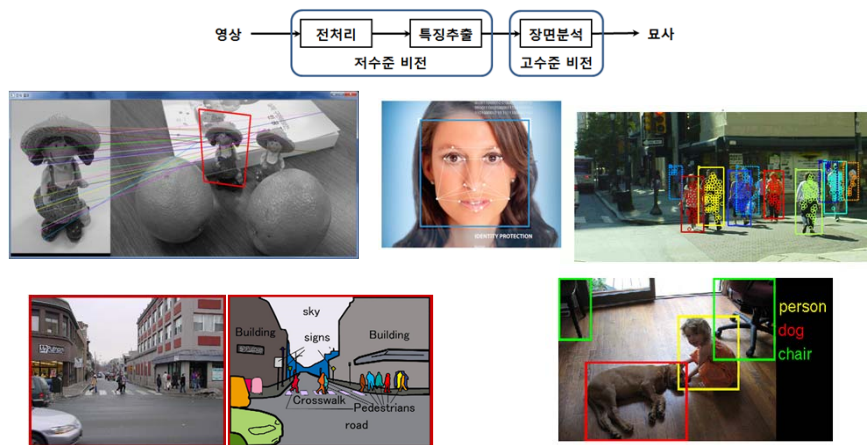
- 사람의 음성 언어를 컴퓨터가 해석해 그 내용을 문자 데이터로 전환하는 처리



주요 응용 분야

❖ 컴퓨터 비전(computer vision)

- 컴퓨터를 이용하여 시각 기능을 갖는 기계장치를 만들려는 분야



<http://www.cis.upenn.edu/~jsji/> <http://www.alternet.org/> <http://www.cs.tau.ac.il/~wolf/OR2/>

주요 응용 분야

❖ 지능 로봇(intelligent robots)

- 로봇틱스(robotics)
 - 로봇에 관련된 기술 분야로서 기계공학, 센서공학, 마이크로 일렉트 로닉스, 인공지능 기술 등을 종합적으로 활용
- 인공지능 기술을 활용하는 로봇

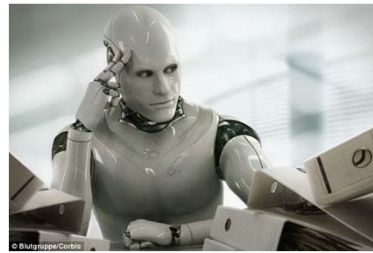
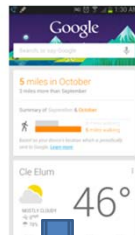


Image : <http://www.dailymail.co.uk/>

5. 최근 동향

❖ 인공지능 비서 서비스

- 애플의 Siri와 구글의 Now
 - 음성으로 대화하면서 필요한 정보를 제공받을 수 있는 서비스
 - 음성 인식 (voice recognition)
 - 자연어처리 (natural language processing)
 - 정보검색 (information retrieval)
 - 추천 (recommendation)
 - 긴급전화 "hey Siri"



최근 동향

❖ 인공지능 비서 서비스 - cont.

▪ 아마존 Echo



- 음성인식 핸드프리(hands-free) 스피커
- **Alexa 음성 서비스** 시스템에 연결
 - 음악 재생, 정보, 뉴스, 스포츠 정보 제공

▪ 마이크로소프트 Cortana



- MS에서 만든 지능형 비서 서비스
- 음성인식, 질의 응답

▪ Viv



- Siri를 개발한 Dag Kittlaus와 Adam Cheyer가 2016.5 데모 공개
- 음성인식, 자연어처리, 다양한 서비스를 연결한 서비스

최근 동향

❖ IBM 왓슨(Watson)

- 자연어로 주어진 질문에 답변을 하는 인공지능 시스템
- 자연어처리, 정보검색, 지식 표현 및 추론, 기계학습 이용 질의에 대한 답변
 - 비구조화된 데이터(unstructured data) 분석
 - 복잡한 질문 이해
 - 답변 제공



Jeopardy! 쇼에 출연하여 퀴즈 경합에서 우승

- **의료, 금융, 유통** 등 다양한 분야에 활용
- 한국어 이해 학습 (2016년) 및 한국시장 진출

최근 동향

❖ 자율주행 자동차(Driverless Car)

- 미국 여러 주에서 무인자동차 시험운행 허가
- 구글 자동차
 - 2009년 이후 2016년 3월 : 2.4백만 km 이상 주행
 - 2016.2.14 최초 자율주행자동차 과실 사고 발생
 - 2020년 출시 예정
- 대부분의 자동차 회사 개발연구 진행
 - Tesla Motors, Volvo, GM, Daimler, Ford, Audi, BMW, Hyundai 등

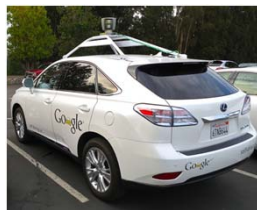


image : www.entrepreneur.com

최근 동향

❖ 로봇틱스(robotics)

- 소셜 로봇 Jibo
 - 대화, 비서, 교육, 엔터테이너
 - 홈자동화



[YOUTUBE / JIBO]

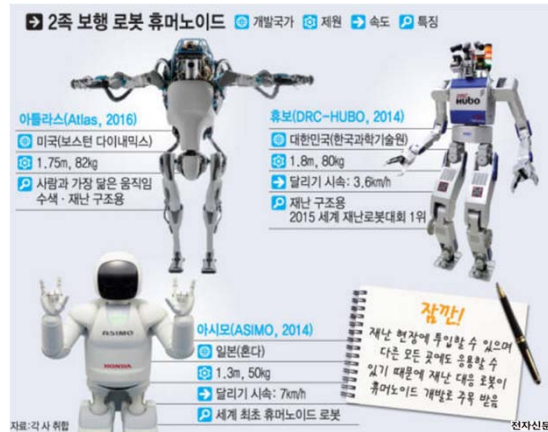
- 캠페니언 로봇 Pepper
 - 일본 SoftBank 감정 인식 대화형 로봇



최근 동향

❖ 로봇틱스(robotics)

▪ 2족 보행 로봇 휴머노이드



최근 동향

❖ 운전 보조 시스템

▪ Mobileye 센서

- 주행 보조 시스템
- 충돌 경고
- 차선 이탈
- 보행자 경고
- 상향등 조절
- 과속 경고



- 자율 주차
- 긴급 제동
- 차선 유지 조향 자동화
- 졸음 방지
- 전방주시 지원

최근 동향

❖ 클라우드 서비스 앱

- 스마트폰을 데이터 입출력 단말기로 사용
- 핵심 처리는 클라우드의 서버에서 담당



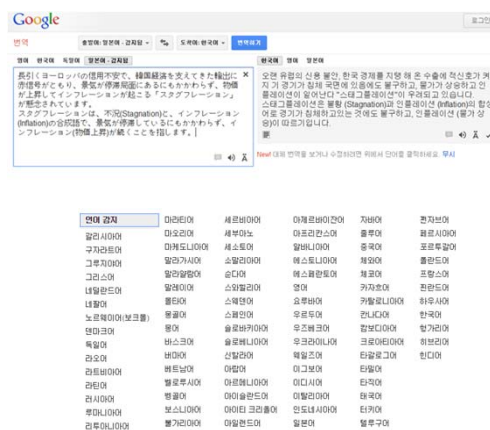
- 구글의 앱 Goggles
 - 이미지 인식 스마트폰 앱
 - 기기에서 찍힌 사진을 기반으로 한 검색

blog.intelligistgroup.com

최근 동향

❖ 기계번역(machine translation)

- Google Translate (translate.google.com)



최근 동향

❖ 마이크로소프트 채팅 봇 테이(Tay)

- 10대 소녀 채팅 봇
 - 미국 18~24세 소셜미디어 사용자 대상
- Twitter에서 2016.3.23 16시간 운영 후 중단
- 인공지능 기술을 통해 **개인화된 대화**(personalized conversation) 학습
 - 테이와 많은 이야기를 할 수록, 대화 상대에 맞춤형 대화를 하게 됨
 - 오도된 학습 사례 발생

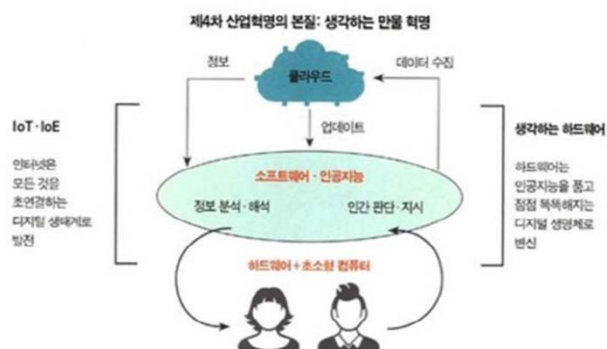


최근 동향

❖ 사이버 물리 시스템 CPS(cyber-physical system)

- 모든 것이 연결되고 보다 지능적인 사회로의 진화
- IoT와 인공지능을 기반으로 사이버 세계와 물리적 세계가 네트워크로 연결된 하나의 통합 시스템

표 2 제4차 산업혁명의 본질



6. 인공지능의 윤리

❖ 인공지능의 윤리

- 마음이 없는 인공지능
- 살상용 자율무기(LAWS: Lethal Autonomous Weapon Systems)
 - 인간의 개입없이 스스로 표적을 찾아내고 제거하는 무기
 - 잘못된 판단의 문제
 - 인명 살상을 위한 프로그래밍



National Robotics Engineering Center of Carnegie Mellon University

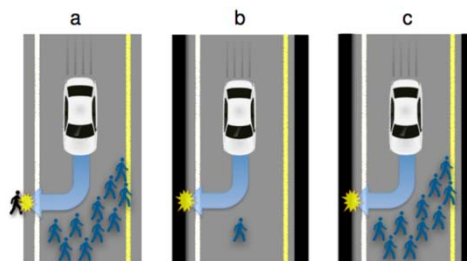


<http://www.briefreport.co.uk/news/taranis-uk-armed-drone-prototype-revealed-2218569.html>

인공지능의 윤리

❖ 인공지능의 윤리 – cont.

- 자율주행 자동차의 돌발 상황에 대한 프로그래밍
 - 모든 가능한 상황에 대한 고려 필요
 - 돌발상황에서 희생자를 선택하는 프로그래밍



Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?
[Bonnefon, Shariff, Rahwan, 2015.10]

인공지능의 윤리

❖ 인공지능의 윤리 - cont.

▪ 로봇 3원칙



Isaac Asimov
(1920-1992)

• 아이작 아시모프 [1942]

1. 로봇은 인간에 해를 가하거나, 혹은 행동을 하지 않음으로써 인간에게 해가 가도록 해서는 안 된다.
2. 로봇은 인간이 내리는 명령들에 복종해야만 하며, 단 이러한 명령들이 첫 번째 법칙에 위배될 때에는 예외로 한다.
3. 로봇은 자신의 존재를 보호해야만 하며, 단 그러한 보호가 첫 번째와 두 번째 법칙에 위배될 때에는 예외로 한다.



▪ 로봇 윤리학 [Lin, Abney, Bekey 2012]

- 로봇공학자의 전문가적 윤리
- 로봇 안에 프로그램된 '윤리 코드'(moral code)
- 로봇에 의해 윤리적 추론이 이루어질 수 있는 자기 인식적 로봇 윤리

▪ OpenAI



Elon Musk
(1971-)

- 2015.12 설립된 비영리 인공지능 회사
- 엘론 머스크(테슬라 모터즈, 스페이스 X, 페이팔)가 10억달러 출연
- 인공지능 발전을 인류의 가장 큰 실존적인 위협으로 간주
- 인류에 혜택이 되는 인공지능 연구

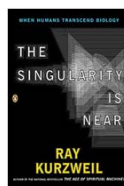
7. 인공지능의 특이점

❖ 특이점(singularity)

- 기술의 수준이 어느 한순간 기하급수적으로 증가하는 시점
- Vernor S. Vinge 교수의 에세이 "The Coming Technological Singularity"에서 사용 (1993)
 - 초인간(superhuman) 인공지능 개발은 인류 종말의 시점일 것

❖ 인공지능의 특이점

- 인공지능이 인간의 지능보다 더 진보하게 되는 시점
- R. Kurzweil의 "The singularity is near"에서 2045년 기술의 특이점 도달 예측



요 약

- ❖ 인공지능의 정의
 - 강한 인공지능, 약한 인공지능
- ❖ 인공지능의 발전과정
- ❖ 요소기술
 - 지식표현, 탐색, 추론, 기계학습, 계획수립, 에이전트
- ❖ 응용 분야
 - 전문가시스템, 자연어처리, 데이터 마이닝, 음성인식, 컴퓨터비전, 지능 로봇
- ❖ 최근 동향
- ❖ 인공지능의 윤리
- ❖ 인공지능의 특이점