인공지능의 미래의 발전의 7단계

- 저자 : Rohit Talwar, Steve Wells, Alexandra Whittington, April Koury, Helen Calle
- 원문 : “The 7 stages of the future evolution of Artificial Intelligence” (2018.6.13)
- 분류 : 일부 내용 정리
- 정리자 : 정백수

인공지능의 미래의 발전의 7단계

1단계	규칙에 기반을 둔 시스템들	사업용 소프트웨어인 RPA(로봇 프로세스 자동화)에서 가정에서 쓰는 도구들, 비행기 자동항해까지, 오늘날 인공지능의 가장 흔한 발현 형태들이다. [주석] 로봇 프로세스 자동화(Robotic process automation)는 소프트웨어를 인공지능 및 기계학습 능력과 함께 사용하여 이전에는 인간이 수행한 덩치가 크고 반복되는 과제들을 다루는 것이다. 이 과제들에는 질의, 계산, 기록과 거래의 유지가 포함된다.
2단계	맥락 이해와 기억	특정 도메인에 대한 일단의 정보를 구축한다. 최고의 인간의 지식과 경험을 기반으로 한다. 지식 베이스는 업데이트될 수 있다. 가장 흔한 형태는 챗봇들, ‘ 로봇 조언자’(roboadvisor)들이다.
3단계	특정 도메인의 전문지식	인간의 능력을 넘어서는 전문지식을 특정 도메인에서 구축. 암 진단과 같은 응용프로그램. 알파고가 가장 유명한 사례이다. 하지만 아직은 하나의 도메인에 국한되며, 알파고 제로(AlphaGo Zero)도 만일 회계감사에서 사기거래를 적발하는 법을 배우기 시작하면 바둑에 대해서 배운 것을 잊어버릴 것이다. [주석] 알파고 제로는 알파고가 이세돌 9단을 이긴 다음 해에 탄생했는데 인간의 도움 없이 스스로의 학습 규칙으로 수천 개의 바둑 시합을 관전했으며 자신만의 전략을 개발했다. 3일 후에 알파고와 싸워서 100대 0으로 이겼다.

4단계	추론하는 기계	이 알고리즘들은 ‘정신 이론’을 가지고 있다. 이는 자신과 타자에게 정신적 상태를 부여하는 능력이다. 가령 믿음, 의도, 지식, 그리고 자신의 논리가 어떻게 작동하는가에 대한 인식을 가지고 있다. 따라서 인간 및 다른 기계들과 함께 추론하고 협상하고 상호작용할 능력을 가지고 있다. 현재 개발 단계인데, 몇 년 후면 상업적으로 응용되는 것을 볼 수 있을 것이다.
5단계	자기인식 시스템(Self Aware Systems) / 인공일반지능(Artificial General Intelligence, AGI)	인간과 같은 지능을 지닌 체계이다. 아직은 존재하지 않는다. 이런 체계를 창조하는 것이 인공지능 분야에서 일하는 많은 사람들의 목표이다. 빠르면 5년 안에 볼 수 있다고 말하는 사람도 있고 결코 이 단계를 성취할 수 없을 것이라고 말하는 사람도 있다. 영화 등에서 알려진 여러 사례들—영화 <2001 스페이스 오디세이>(2001 A Space Odyssey)의 우주선 컴퓨터 할(HAL)에서 텔레비전 시리즈 Humans의 ‘신스(Synths)까지—이 있다.
6단계	인공초지능 (Artificial SuperIntelligence, ASI)	모든 도메인에서 가장 똑똑한 인간의 수행 능력을 능가한다. 사실 그 능력이 어떻게 인간의 지능을 능가하는지 분명하게 정식화하기 힘들다. 그러나 기아나 기후변화 같은 현재의 세계가 처한 문제를 ASI가 푸는 것을 상상해볼 수 있다. ASI는 또한 새로운 과학 분야들을 창조하고 경제체계를 재설계할 수 있으며 전적으로 새로운 양태의 거버넌스 모델을 발전시킬 수도 있다. 언제 이것이 가능한지에 대해서도 전문가들의 견해가 갈린다. 다음 10년 안에 볼 수 있으리라고 생각하는 전문가는 거의 없다. 영화 Her나 Ex Machina가 우리의 테크놀로지가 우리를 능가할지도 모를 세계에서 ASI의 가능성을 흥미롭게 묘사한다.

7단계	특이점(Singularity)과 초월(Transcendence)	ASI가 가능하게 하는 기하급수적 발전경로가 인간 능력의 대대적인 확장을 가져올 수 있을 것이라는 생각이다. 언젠가 우리의 뇌를 서로 연결할 수 있으며 현재의 인터넷의 미래의 계승자와 연결할 수도 있을지 모른다. 이 ‘하이브 마인드(hive mind)가 우리로 하여금 아이디어들을 공유하고 문제를 집단적으로 해결하며 서로의 꿈에도 관찰자로서 혹은 참여자로서 접근할 수 있게 해줄 것이다. 생각을 한 단계 더 끌고 나가면 우리는 또한 인간의 신체의 한계를 뛰어넘을 수 있으며 동물. 식물, 날씨체계, 자연환경과도 연결될 수 있다. 구글의 엔지니어링 이사인 커즈와일(Ray Kurzweil) 같은 일부 특이점 주창자들은 2045년쯤에는 광범한 과학 및 테크놀로지 분야들을 가로지르는 기하급수적 진보의 결과로 특이점이 발생하는 것을 볼 수 있다고 한다. 다른 이들은 그것은 결코 불가능하며 우리는 결코 인간의 의식을 포획하거나 디지털화할 수 없을 것이라고 주장한다. [주석] ‘특이점(singularity)이란 일반적으로 어떤 기준을 상정했을 때, 그 기준이 적용되지 않는 점을 이르는 용어로 여러 분야에서 사용된다. 물리학의 경우 빅뱅 이전에 밀도와 질량이 무한대가 되는 점을 가리킨다. 여기서는 ‘기술적 특이점(technological singularity)을 줄인 말로서 ASI가 일으킨 가속화된 기술 발전의 결과로 인간 문명에 측량할 수 없는 변화가 일어나는 점을 가리킨다.