

인공지능특강 독후감 Report

[인공지능, 컴퓨터가 인간을 넘어설 수 있을까?]



제출일	2017.06.14 (수)
학과	컴퓨터공학과
학번	1731036007
이름	이 세 한
담당교수	한연희 교수님

내가 중학생이던 10년 전까지만 해도 인공지능은 SF 소설이나 영화에서만 접했던 기술이지만 현재는 제4차 산업혁명으로 인해 앞으로 우리의 삶에 있어 빠질 수 없는 현실로 다가오고 있다.

2016년 3월 9일. 구글의 알파고라는 인공지능 시스템과 한국의 바둑기사인 이세돌 사범의 바둑 대결이 시작되었고, 총 5일 동안의 대국에 걸쳐 알파고는 4:1로 승리를 거머쥐었다. 바둑 기사분들 뿐만 아니라 기자, 그리고 그 대국을 지켜보던 많은 사람들은 충격을 받을 수밖에 없었다. 인간이 인공지능에게 쉽게 패배를 한 것이기 때문이다.

나 또한 이 대국을 지켜보면서 공포감이 생기기 시작했다. 아직 인공지능은 계속 연구를 하고 있는 단계이긴 하지만 시간이 지난 뒤 인간과 똑같이 행동하고 생각할 수 있는 인공지능이 나타난다면 과연 우리 인간들은 어떤 상황에 처하게 될 것인가?

이번에 읽은 책에서는 이러한 인공지능에 대해 설명하기 위해 ‘로봇’을 예로 들고 있다. 한 파트를 할애하여 설명할 정도로 로봇도 중요한 이야기지만 로봇 못지않게 중요하다고 생각하는 것은 바로 ‘스스로 생각하는 인공지능 시스템’이다. 그리고 ‘스스로 생각하는’이라는 것으로 인해 인간은 인공지능의 윤리적 문제에 대해 끊임없는 이야기를 할 것이다.

“로봇이 가사도우미가 되는 세상이 곧 닥친다고 해보자. 오전 11시가 다가오는 시간, 아빠는 거실에 있는 로봇한테 리모컨을 가져오게 해서 아침 뉴스를 시청 중이다. 그런데 엄마도 다른 채널을 보고 싶어서 로봇더러 리모컨을 가져오라고 한다. 로봇은 리모컨을 엄마에게 가져다주기로 한다. 아빠는 당황한다. 그런데 로봇이 아빠에게 당신은 어제 좋아하는 아침 뉴스쇼를 봤으니 이번에는 엄마가 원하는 TV 프로를 보는 것이 합당하다고 설명한다. 이런 상황은 일상에서 흔히 일어날 만한 윤리적 결정에 관한 문제가 어떤 곤란함을 일으킬지 보여준다. 기계가 이런 일을 매끄럽게 다루기는 어려울 것이다. ... 예상치 못한 새로운 상황에 로봇 스스로 윤리적 원칙을 적용하게 하는 것이 해결책이 된다. 리모컨을 누구에게 먼저 줄지 결정하기보다는 누구에게 새 책을 먼저 읽으라고 권할지를 고려해보는 것이다. 이런 접근 방법을 쓰면 자신의 행동을 판단해야 하는 순간에 로봇이 윤리적 원칙에 따라 움직이도록 하는 장점도 있다. 로봇의 행동은 기본적으로 인간이 느끼기에 편안해야 한다는 점을 생각하면 이는 매우 중요한 부분이다. 덧붙여 말하면 윤리적 로봇을 만드는 일은 철학자들이 실생활을 들여다보게 함으로써 윤리 자체를 발전시키는 효과도 있다. 터프츠대학 철학자 대니얼 데넷(Daniel C. Dennett)은 이를 [인공지능이 철학을 정직하게 만든다.]라고 표현했다.” (책의 본문 중에서...)

위 본문의 내용은 로봇에 탑재된 인공지능 시스템은 인간의 윤리적 사상에 접근을 하여 스스로 생각한 뒤 판단하는 내용을 담고 있다.

로봇의 행동에는 어쩔 수 없이 윤리적 원칙이 적용이 되어야 한다. 인간을 도와주

는 행동이라고 해도 모든 행동이 윤리적인 행동일 수는 없는 것이기 때문이다. 앞으로 개발될 모든 인공지능 로봇에는 반드시 인간의 윤리적 내용이 프로그래밍되어 개발되어야 한다. 이를 통해 자연스럽게 인간의 윤리 철학에 대해 다시 한 번 생각해볼 수 있는 계기가 마련될 것이다.

즉 인공지능의 발전이 인간의 옳고 그름에 대해 다시 한 번 논의해보는 밑거름이 되는 셈이다.

이 책을 읽으면서 생각난 것은 유명한 SF 작가인 ‘아이작 아시모프(Isaac Asimov)’의 작품인 [I Robot]이었다.

[I Robot]에서는 인공지능 로봇의 윤리적 판단을 위해 “로봇의 3원칙”이라는 것이 있다.

- ① 로봇은 인간에게 해를 끼쳐서는 안 되며, 위험에 처해 있는 인간을 방관해서도 안 된다.
- ② 로봇은 인간의 명령이 1원칙에 어긋나지 않는 한 반드시 인간의 명령에 복종해야 한다.
- ③ 로봇은 1원칙과 2원칙에 어긋나지 않는 한 자기 자신을 보호해야 한다.

이 내용은 로봇을 대상으로 한 내용이지만 사실 ‘스스로 생각하는 인공지능 시스템’에 대한 내용이라고 이해해도 무방하다는 생각이 든다.

위 3원칙은 서로 모순적인 내용을 담고 있다. 이는 인간의 이기주의적인 생각이 담겨져 있기 때문이다. 아직 인간도 도덕적으로 완벽하지 않고, 윤리적으로도 완벽하지 않은데, 과연 인공지능에게 인간의 도덕적 윤리 철학을 가르칠 수 있을까?

책에는 다음과 같은 이야기도 나온다.

“흥미로운 사실은, 로봇 윤리가 실제 윤리학 분야에도 영향을 미친다는 점이다. ‘현실 세계’에서의 인공지능을 연구하다 보면 단지 윤리학의 학문적 결과를 기계에 적용하는 데서 끝나지 않고, 인간이 행하는 윤리적 행동이란 무엇인지 더 깊이 생각해보게 되기 때문이다. 적절하게 훈련된 기계는 항상 공정한 판단을 하기 때문에 대부분의 인간보다 훨씬 더 윤리적으로 행동한다. 인간이 항상 공정한 판단을 하기란 어렵다. 윤리적으로 움직이는 로봇이 우리에게 더 윤리적으로 행동해야 한다는 생각을 갖게 해줄지도 모를 일이다.” (책의 본문 중에서...)

이 내용은 나에게 많은 생각을 던져 주었다. 공정한 판단? 어떤 것이 공정한 판단일까? 인간의 도덕적 윤리 철학에 알맞으면 그것이 무조건 공정한 판단일까? 예를 들어 이러한 상황이 발생하면 어떨까.

하나의 기차길이 두 개의 길로 갈라지는 길이 있다고 가정 하자. 현재 기차는 브레

이크가 고장난 상태로 달려오는 중이며, 두 개의 길 중 한 곳으로는 무조건 달려가야만 탈선사고가 일어나지 않을 수 있다. 이 때 한 쪽 길에서는 어린 아이 2명이 장난을 치면서 놀고 있고, 다른 한 쪽 길에서는 철길 안전점검을 위해 안전관리 담당자 5명과 철길 보수공사를 위한 인부 10명이 있다. 어린 아이 2명이 재밌게 놀고 있는 길은 길에는 이상이 없지만 철길 보수공사를 하고 있는 길은 공사 중이라 잘못하면 기차 내의 승객들도 피해를 입을 수 있는 상황이다. 기차는 어떻게든 하나의 길을 선택해서 가야하는 상황이고 이 기차를 운전하는 것은 다름 아닌 인공지능 로봇이다. 이 때 어느 쪽 길로 가야만 할 것인가? 어린 아이를 살리기 위해 공사 현장으로 가야할까? 아니면 인명피해를 줄이기 위해 어린 아이들이 있는 곳으로 가야 할까? 어느 것이 공정한 판단인가?

그렇다. 이는 美 하버드대학교의 정치철학 교수인 ‘마이클 샌델(Michael J. Sandel)’의 [정의란 무엇인가?]에 나오는 내용과 비슷하다.

Justice. 인간의 생활 속에서도 이처럼 공정한 판단이 제대로 이루어지기 힘든 상황이 비일비재하다.

물론 인공지능은 여러 가지 계산을 통해서 인간보다 더 윤리적인 선택을 할 수도 있지만(예를 들어 법정에서의 판정) 그 한계는 분명 존재할 것이다.

앞으로 인류는 인공지능을 개발하면서 이러한 한계들을 어떻게 극복할 것인지를 깊게 고민해야 할 것이다.