

인공지능(人工知能)의 시대, 전통문화연구회(傳統文化研究會)의 과제



金 炫

본회 부회장

한국학중앙연구원 인문정보학 교수

I.

대화형 인공지능 Chat GPT(Generative Pre-trained Transformer)에 대한 논의로 세계가 떠들썩하다. 작년 12월에 출시된 이후 그로부터 반년도 지나지 않아 지난 5월 3일에는 학습 데이터량을 획기적으로 증대시킨 새 버전이 발표되었다. 여기에서는 한국어를 비롯한 다국어(多國語) 처리 기능 향상이 특별히 눈에 띈다.

도대체 이 소란이 무엇인지 이해가 안 되는 사람도 많겠지만, IT 산업계(產業界)는 물론이고 학계(學界)와 교육계(教育界)도 이것으로부터 받은 충격이 이만저만하지 않다.

오늘날의 인공지능 프로그램이 우리의 관심사인 “한문(漢文) 공부(工夫)와 무슨 관계가 있는지 이야기하기 위해 문제의 Chat GPT의 간단한 사용 예시를 보이도록 하겠다.¹⁾

나의 질문 :

“子曰學而時習之, 不亦說乎!” 이 글에서 ‘학(學)’과 ‘습(習)’은 무슨 뜻이고 어떠한 차이가 있습니까?

Chat GPT의 대답 :

이 문장은 고대 중국의 철학자인 공자(孔子)의 말씀입니다.

“子曰學而時習之, 不亦說乎!” 라는 문장에서 ‘학(學)’은 ‘배우다’라는 뜻이며, ‘습(習)’은 ‘습득하다’라는 뜻입니다. 이 둘은 비슷한 의미를 가지고 있지만, ‘학(學)’은 지식을 습득하는 과정 자체를 가리키고, ‘습(習)’은 습득한 지식을 꾸준히 복습하고 익히는 것을 말합니다.

즉, 이 문장에서 ‘학이시습지(學而時習之)’는 “지식을 배우고 꾸준히 복습하여 습득하라”는 뜻입니다. 이것이 공자의 교육 철학 중 하나인 “학습의 반복성”을 강조하는 말씀입니다. 이 말은 공자가 학습에 대해 단순한 지식 습득을 넘어서 꾸준한 복습과 연습이 중요하다는 것을 강조하고 있습니다.

1) Chat GPT와의 대화할 수 있는 사이트의 주소
<https://chat.openai.com/>

나의 질문 :

공자의 이 말씀에 대해 조선의 유학자 다산 정약용은 “學所以知也. 習所以行也. 後世之學, 學而不習, 所以無可悅也.” (정약용(丁若鏞), 《논어고금주(論語古今註)》) 라고 풀이하였습니다. 이 문장을 한국어로 번역하고 정약용의 주장은 무엇인지 설명해 주세요.

Chat GPT의 대답

다산 정약용은 "공부는 지식을 얻기 위한 것이며, 연습은 실천하기 위한 것이다. 그래서 나중 사람들이 공부할 때, 공부만 하고 연습을 하지 않으면 기쁨을 느낄 수 없다." 라고 번역할 수 있습니다.

정약용은 이 말씀을 통해 지식을 얻는 것이 중요하지만, 지식만 갖고 있어서는 실제로 그것을 활용하는 데에는 한계가 있다는 것을 말하고 있습니다. 따라서, 얻은 지식을 실제로 활용하기 위해서는 연습과 실행이 필요하다는 것을 강조하고 있습니다. 즉, 공부와 실천이 모두 중요하다는 것을 말하고 있는 것입니다.

Chat GPT가 어느 수준의 한문 문장을 번역할 수 있고, 그 내용을 풀어서 설명할 수 있는지 알아보려고, 여러 가지 문헌의 문장을 가지고 이런 식의 질문을 던져 보았다. 한국어 번역은 전체적으로 어색했지만, 영어로 번역하라고 하면 놀라우리만큼 정확한 답을 주는 경우가 많았다. 영문 번역이 존재하는 중국 고전은 당연히 인공지능이 그것을 학습했을 터이니 정확한 번역이 가능할 것이다. 그러나 영어 번역이 있을 것 같지 않은 우리나라의 한문 텍스트도 꽤 정확하게 번역하는

것은 이 인공지능이 상당한 응용 능력이 있음을 알게 한다. 이것은 한국한문이나 중국 한문이나 “한문”으로서의 공통성이 있는 것이고, 인공지능이 그 공통성에 기초해서 유의미한 예측을 할 수 있게 된 때문이다. 여유당전서(與猶堂全書)나 조선왕조실록(朝鮮王朝實錄)의 문장 하나하나는 중국 고전에 없는 표현일 수 있어도 그런 식의 표현을 만들어 내는 용어와 용법은 중국 문헌에도 적지 않다는 것이다. 물론, 한국의 고유한 제도나 개념어 등 ‘공통성’에서 벗어난 단어가 들어간 텍스트의 경우, 그 부분은 틀리거나 이상한 번역으로 표현되기도 한다.

현재의 Chat GPT는 고전 한문 번역을 위해 특별히 훈련된 AI가 아니다. 전지구의 인터넷 서버에 존재하는 방대한 양의 데이터(500TB, 책 3천만권이 넘는 분량)를 망라적으로 학습하는 가운데 고전 한문에 관한 지식이라고 할 수 있는 것도 더불어 얻게 된 것으로 보인다. 아직까지는 미흡한 점이 많다 하더라도 그 능력이 점점 더 증대되어 갈 것이고, 결국 동양고전(東洋古典)의 세계도 ‘인공지능’과의 동행하는 길을 찾지 않고는 더 이상 존립할 수 없는 시대에 이를 것이다. 그리 머지 않아서.....

II.

그동안 한문 고전의 교육과 연구가 주된

관심사였던 전통문화연구회의 가족들도 이 제는 이 ‘인공지능과의 동행’에 관해 보다 분명한 입장 정리와 실천이 있어야 할 것이다. 무엇부터 시작해야 할까?

우선, 미래 세대가 동양 고전을 공부하는 방식은 우리가 해 온 공부법과는 완전히 달라질 것이라는 사실을 인정하고, 그것에 대비하는 새로운 교육 프로그램을 준비해야 한다. 디지털 원어민이라고 해야 할 젊은 세대들은 새로운 한문 문장을 만나면 자신이 터득한 문리로써 그것을 해석하거나 선생에게 묻기보다는 인공지능에게 먼저 그 해석을 구할 것이다. 미래의 한문 선생은 모든 것을 스스로 가르치려 하기보다는 인공지능을 조력자(助力者)로 활용할 수 있어야 한다.

상업적인 현대 외국어 번역 시장에서는 이미 이런 식의 해법이 낫설지 않다. 영어 번역의 경우, 전문 번역가는 원고를 처음부터 직접 번역하기보다 인공지능(기계번역 프로그램)에게 초벌 번역을 맡기고, 자신은 그 결과 속의 오류를 수정하거나 그 뜻이 사람들에게 잘 읽혀질 수 있도록 수정하는 일을 전문적으로 수행한다. 이런 식의 일처리가 시간과 비용을 절약하는 방법이라고 인식하는 번역 회사들이 점점 늘어나고 있다. 하지만 이러한 변화에 대해 비용 절감을 위해 사람의 일을 기계가 대신하게 하는 것이라고만 이해해서는 안 된다. 인공지능의 번역결과를 검토

하고 그 속의 오류를 수정하는 “새로운 전문성”이 강도 높게 요구되는 일이다. 예전의 기계번역 결과물은 그 표현이 투박했다. 그래서 문장을 매끄럽게 만드는 윤문이 사람의 손으로 하는 후처리 작업이라고 인식하곤 했다. 그러나 요즘 Chat GPT가 보이는 상황은 그렇지 않다. 적어도 영어문장의 경우, 그 표현이 상당히 자연스럽다는 것이 중론이다. 그러나 그 매끄러운 문장 속에 아는 사람이 아니면 속아 넘어갈 틀린 이야기가 섞여 있는 경우가 종종 발견되곤 한다. 인공지능이 만든 결과물을 검토할 전문가는 언어를 다듬는 윤문가가 아니라 그 내용의 사실성을 검증할 수 있는 해당 분야의 지식 전문가여야 한다.

미래의 한문 선생은 학생들에게 한문 구문의 분석 방법을 가르칠 필요가 없을지도 모른다. 학생들은 이미 훨씬 쉽게 접근할 수 있는 인공지능으로부터 한문 텍스트에 대응하는 영어나 한국어 대역문(對譯文)을 얻을 수 있고, 또 그 내용을 친절하게 설명해 주는 해설문(解說文)도 얻을 수 있다. 선생의 역할은 학생들로 하여금 인공지능의 준 대답 중에서 신뢰할 수 있는 것과 그렇지 않은 것을 가려낼 수 있게 하는 안목을 갖게 하는 것, 그리고 인공지능이 도울 수 있는 것의 한계를 넘어서는 새로운 지식의 탐구 방법을 알려 주는 것이다. 인공지능 시대의 한문 선생

은 동양고전의 지식세계에서 깊은 통찰력과 구체적인 정보력을 지닌 지식전문가(知識專門家)여야 한다.

손에서 휴대전화를 놓고서는 살 수 없는 현대인들은 이미 일상생활에서 인공지능과 동행하는 존재가 되어버렸다. 정보 검색창에서 찾아지는 검색 결과 하나도 포털이 운영하는 인공지능 시스템에 의해 걸러진 것이다. 유튜브가 알아서 보여주는 연주 영상과 교양 프로, 온라인 쇼핑몰에서 구매를 권유하는 물품 목록이 모두 나의 취향을 데이터화한 인공지능의 선택이다.

동양고전(東洋古典)의 교육 현장에서도 ‘인공지능’과의 동행은 불가피하다. ‘인공지능이 한문을 해석하고 설명해 주니, 더 이상 동양고전 교육은 불필요하다’는 생각은 ‘동행’의 필요성과 방법을 알지 못하는 이들의 푸념일 뿐이다. 미래의 한문 선생은 ‘주제넘게 한문 까지도 해석하는’ 인공지능을 무시하거나 적대시하기보다는 그것이 학생들의 고전 공부에 도움을 줄 수 있다는 사실을 인정하고, 인공지능의 교육적 기능이 더 유효하고 건설해질 수 있도록 하는 역할을 해야 한다.

III.

Chat GPT와 같은 대화형 언어 변환기가 대변하는 오늘날의 인공지능은 처음 보는 한문 문장도 번역해낼 뿐 아니라, 그 뜻이 무엇

인지 부연 설명해 줄 수 있는 수준으로 진화했다. 그것은 GPT라는 인공지능 엔진이 지난 수년간 우리의 상상을 뛰어넘는 방대한 분량의 텍스트를 학습하면서 스스로 터득한 ‘말주변’이다.

내가 Chat GPT의 놀라운 능력에 감탄하면서도 그 능력을 ‘말주변’이라고 폄하하는 이유는 그것이 주는 정보는 신뢰할 수 있는 ‘근거’로 ‘검증’할 수 있는 지식이 아니기 때문이다. Chat GPT가 둘러대는 이야기들이 그럴 듯하게 들려도, 그 안에는 비슷하지만 틀리거나 관계가 없는 이야기도 적지 않다. Chat GPT는 기본적으로 무엇인가 실마리가 주어졌을 때 그것과 연관하면서 말이 되는 문장을 만들어내는 데 특화된 기계이다. 그러다보니 15세기의 역사적 사실을 이야기 하다가 유사한 용어로 설명되는 19세기의 일을 갖다 붙이는 식의 사고도 일어난다. 역사적인 인물 여러 사람의 이름이 거론되는 한문 문장을 번역하라고 시키면, 그 문장 자체는 하자 없는 영문이나 한국어 현대문으로 번역한다. 그리고 나서 그 사람들이 누군가고 물으면 맞는 이야기와 무관한 이야기를 섞어서 누군지 모를 가공의 인물을 만들어내기도 한다.

사실상 지식 동반자로서의 인공지능은 아직까지 극복해야 할 많은 과제를 안고 있다. Chat GPT의 가장 심각한 문제는 틀린 이야기를 하면서도 스스로 그것이 틀렸다는 것을

모른다는 것이다. 하지만 여기서 맞거나 틀리다고 하는 것은 질문의 의도를 알고 있는 인간의 관점에서 하는 이야기이지 인공지능에게 그 책임을 물을 수 있는 것이 아니다. 인공지능의 모든 대답은 자신이 학습한 데이터를 가지고 스스로 만들어낸 사고의 논리에 따라 충실하게 답변한 것이라고 할 수 있다. 만일 어떤 특별한 지식 영역 안에서 ‘옳다’고 인정되는 답을 인공지능이 그대로, 또는 가깝게 찾아주기를 원한다면, 우리는 그 분야의 정리된 지식을 집중적으로 교육하는 방식으로 인공지능의 ‘전문성’을 제고시킬 수 있다. 이런 일을 ‘지도학습’(supervised learning)이라고 한다. ‘지도 학습’보다는 느슨하게 ‘주어진 지식과 인공지능 스스로 습득한 지적능력을 균형적으로 융합할 수 있게 하는 것을 ‘강화학습’(reinforcement learning)이라고 한다. 현재 세계인들에게 선보인 Chat GPT는 언어적인 관점에서 ‘자연스럽게’ 대답하는 능력을 이 강화학습의 과정을 통해 습득했다.

인공지능이 오답을 주기도 한다고 해서 그것을 외면하는 것은 길이 아님을 우리는 알고 있다. 그렇다고 해서 오류가 계속 확대 재생산되는 것을 방관해서도 안 된다. 인공지능이 동양고전 교육과 연구의 동반자가 될 수 있게 하기 위해서는 그것이 이 세계의 지식에 대해서도 보다 건설한 지식과 지능을 쌓을 수 있도록 도와야 한다. 그 방법은 동양

고전 연구자 스스로 인공지능에게 동양고전 지식을 전수하는 ‘지도 학습’이나 ‘강화 학습’에 참여하는 것이다. 인공지능 관련 기술의 특징은 그것을 어느 특정 조직이 폐쇄적이거나 독점적으로 운용하기보다는 다양한 외부 조직에서 그 기반 기술을 응용할 수 있도록 허용하고 협업을 촉진하는 문화 속에서 발전하고 있다는 점이다. GPT의 경우도 마찬가지로 GPT API(Application Programming Interface)를 제공함으로써 특정 목적으로 활용될 수 있는 응용 시스템 개발의 문을 열어두고 있다. 이것을 이용하면, 우리 스스로 동양고전의 지식세계에서 보다 의미 있고 정확한 지식을 전달할 수 있는 특수 목적으로 GPT 응용 프로그램을 개발하는 것도 가능하다. 물론 이를 위해서는 동양고전 데이터의 특수성에 맞춰 GPT API의 기본 기능을 보완하는 맞춤형 알고리즘을 개발해야 하고 또 ‘지도 학습’이나 ‘강화 학습’에서 충분한 효과를 거둘 수 있도록 충분한 규모의 학습 데이터를 준비해야 한다.

동양고전의 지식 세계에서 유의미한 학술적 지식의 증계자 역할을 할 인공지능 응용 시스템의 개발은 전통문화연구회의 새로운 미래과제이다. 이제 우리는 이를 위한 실천 전략과 구체적인 로드맵을 마련해야 할 시점에서 있다. ㉞