새로운 코드 학습 시스템

KAIST 전산학부 유제인

AI 기술, 특히 Copilot과 ChatGPT와 같은 도구들은 프로그래밍의 많은 부분을 자동화하고 있다. 이로 인해 개발자들은 반복적이고 단순한 작업에서 벗어나 보다 고차원적인 작업에 집중할 수 있게 되었다. 그러나 현재의 AI는 통계적 언어 모델에 기반하고 있어, 복잡하거나 익숙하지 않은 문제 상황에서는 부정확하거나 불완전한 결과를 생성한다. 특히 보안과 같이 높은 정확성과 신뢰성이 요구되는 분야에서는 이러한 한계가 더욱 두드러진다. 이에 따라 장기적으로는 자연어처리가 아닌 논리 기반의 AI가 프로그래밍 문제 해결에 더 적합하다고 주장한다.

ChatGPT 없으면 과제 어떻게 했을까? Copilot 없었으면 이거 못했을 거야! 주변에서 흔히 들을 수 있는 말이다. AI는 우리 삶에 깊이 침투해있고, 점점 더 발전 속도가 빨라지고 있기 때문에 AI 이전의 삶을 상상하기 어렵다. 특히 나같은 전산학도들에게는 AI를 이용한 프로그래밍이 점점 익숙해지고 있다. Copilot이 단순한 코드를 너무 잘 작성해주기 때문에 Copilot 이전의 삶이 상상도 되지 않는다. 이렇게 지루한 코드를 일일히 작성해야 했다니!

하지만 당연하게도 AI는 모든 것을 해결해주는 마법이 아니다. 단순한 코드는 잘 작성해주지만, 조금만 복잡하거나 생소해져도 제대로 된 코드를 작성하지 못한다. 예를 들어, OCaml을 이용한 정보보호개론 과제를 해결할 때에는 거의 도움이 되지 않는다. AI는 기존의 코드를 학습해 비슷한 패턴으로 따라하기 때문에 학습 데이터가 불충분한 경우에는 제대로 된 코드를 작성할 수 있다. 이는 현재 자연어처리 기술의 특성이기 때문에 단지 프로그래밍 분야 뿐만 아니라 다른 모든 분야에도 똑같이 적용된다. 하여튼, 이러한 오류 때문에 AI가 쓴 코드를 그대로 쓸 수는 없다. [AI, 어떻게 해야 신뢰할 수 있을까?] 글에 나온대로 주객이 전도되어 AI가 짠 코드를 하나하나 체크해주어야 한다.

이런 오류는 내게 크게 심각하게 다가오지 않았다. 나는 AI를 사용해 프로그래밍해봤자 수업 과제에 불과했기 때문이다. Copilot이 쓴 코드를 확인하고, 오류가 있으면 수정하고, 아예 못쓸만하면 아직 발전하지 못한 AI 기술을 탓하며 내가 작성하면 그만이었다. 하지만 보안이 연결된다면 이런 오류는 심각한 문제와 직결된다.

글에서는 이 문제에 대한 해결방법 중 하나로 새로운 코드 학습 시스템을 제안하고 있다. 통계적인 개연성만을 고려하는 기존의 자연어처리 모델을 이용하는 것이 아니라 오류 분석기의 결과를 학습 데이터로 이용하는 것이다. 좋은 접근 방법이라고 생각한다. 지도 학습은 이미 너무나도 연구가 많이 되어있고, 글에서 언급했듯이 소프트웨어의 오류는 정의하기 쉽기 때문이다. 기존의 자연어 처리 모델과 지도 학습 기술을 이용해서 구현할 수 있을 것이라 생각한다.

하지만 나는 조금 더 장기적이고 근본적인 방향의 변화를 바란다. 지금의 자연어처리 모델은 본질적으로 확률적이다. 그 결과 AI는 '그럴듯한' 코드를 생성하지만, '논리적으로 올바른' 코드는 보장하지 않는다. 자연어는 논리보다는 문맥에 따라 의미가 결정되지만, 프로그래밍은 철저하게 논리에 기반한 활동이다. 그럴듯한 말로 대화는 할 수 있어도, 프로그래밍은 그럴듯함만으로는 안 된다. 따라서 프로그래밍에 적합한 AI는 자연어 모델이 아니라, 논리 기반 AI여야 한다.

SW 보안의 문제도 마찬가지다. 예를 들어, 접근 제어 문제를 생각해보자. 이 프로그램에 권한 문제가 있는지 여부는 논리적으로 판단할 수 있다. 접근 제어 행렬에서 권한을 따져가면서 문제 여부를 논리적으로 계산할 수 있다. 그렇게 작성한 것이 지난 과제에서 구현한 monitor와 analyzer 아닌가.

그렇다면 논리 기반 AI는 어떻게 설계해야 할까? 이 질문은 과목의 범위를 넘어가므로 구체적으로 다룰 수는 없지만, 내가 상상하는 바는 자연어를 술어논리로 번역하고, 이를 SAT Solver처럼 논리식 계산으로 해결하는 방식이다. 다소 원시적이고 복잡해 보일 수 있지만, 통계에 기반한 현재의 자연어처리 모델보다는 논리적인 문제 해결에 훨씬 적합할 것이다.