

Generative Chemical Transformer: Attention makes neural machine learn molecular structures via text

김현승, 나종걸¹, 이원보[†]
서울대학교; ¹이화여자대학교
(wblee@snu.ac.kr[†])

목적에 부합하는 물질을 발굴하고, 이를 대량 생산할 수 있는 공정을 설계하는 것은 인간의 삶의 질을 향상시키는 데 있어 중요한 문제다. 본 연구에서는 인공지능을 활용한 소재 개발 가속화의 일환으로, 화학언어를 학습한 인공신경망을 활용해 목표하는 다중 조건들을 동시에 만족할 수 있는 분자구조 역설계 방법론을 제안한다 (Generative Chemical Transformer, GCT). GCT는 conditional variational generator에 Transformer의 언어 인지 능력을 내재시킨 형태다. GCT의 attention mechanism은 line notation로 인해 의미의 불연속성이 발생하는 화학 언어의 한계를 넘어 분자 구조에 대한 깊은 이해를 가능하도록 한다. GCT에 의해 생성된 분자의 품질은 정량적으로 평가되었다. 화학언어의 문법 및 화학규칙을 잘 만족함과 동시에 실제 분자와 유사한 패턴을 보이며, 주어진 다중 조건들을 동시에 만족하는 것이 확인되었다. 또한 학습 데이터셋에 존재하지 않는 새로운 분자구조를 역설계 할 수 있었으며, 분자 당 수백 milliseconds 이내에 역설계가 가능한 것이 확인되었다. 이러한 인공지능 기반 분자구조 역설계 기법은 소재 개발 과정을 가속화하여 인간의 삶의 질을 향상시키는 데 기여할 것으로 기대된다.