FSE 2022 Trip Report

변지석, 고려대학교 소프트웨어 분석 연구실
Singapore, 2022 November 13 ~ 18

들어가며

소프트웨어 공학 분야 최고 수준 학회들 중 하나인 FSE (Foundations of Software Engineering) 에 참석했다. 코로나로 인해 20 년, 21 년에는 오프라인 발표가 없다가, 올해 들어 오프라인 발표 기회가 생겨 차수영 교수님, 송도원, 오원석 셋의 발표길에 함께 갈 기회를 얻었다. 덕분에 2022 년 11 월 13 일부터 11 월 18 일까지 학회 참석 일정을 연구실 사람들과 함께했다.

이 보고서를 통해 학회에서 경험하고 생각했던 내용들과 인상깊게 들었던 발표들, 그리고 싱가포르에서 느낀 점들을 공유하려 한다.

FSE 2022

세션 진행

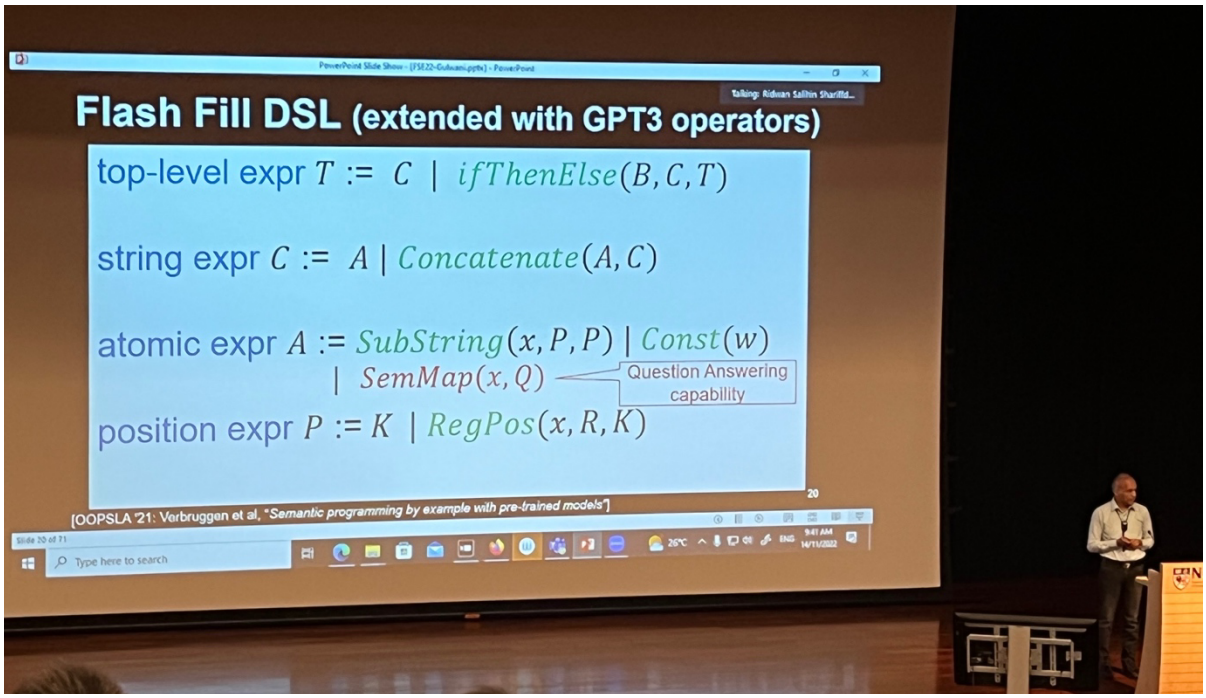
학회는 아침 9 시부터 오후 5 시 반까지 국립 싱가포르 대학 UTown 의 Stephen Riady Centre 의 여섯 강의실에서 진행되었다. 수백명을 수용할 수 있는 대강당과, 150 명 수용규모의 네모난 계단식 강의실 4 곳, 그리고 그 절반크기의 반원형 계단식 강의실이었다. 한 시간대에 최대 여섯 세션이 동시에 진행되었는데, 지난 20, 21 년에 붙은 논문들도 올해 발표기회를 주었기 때문에 발표 수가 크게 늘어났기 때문이었다. UTown 은 커다란 타원형 잔디밭을 건물들이 둘러싼 모습이었는데, 강의실이 위치한 건물은 건물의 커다란 입출구에 문을 따로 두지 않아 바람이 잘 통하고 개방된 느낌을 주었다. 대강당 앞뒤쪽으로 식사와 간식을 동일하게 제공하는 자리가 두 곳 있었고, 사람들은 그 주변 좌식 테이블과 조그만 원형 입식 테이블에 기대어 얘기를 나누었다. 덥고 습한 기후에 맞춰 모든 강의실은 항상 에어컨을 켜 두고 있었는데, 싸늘한 에어컨 바람에 추워하면서도 인상깊게 들었던 발표들을 소개한다.

Stephan Riady Centre 바깥. 계단을 따라 올라가면 문 없이 안으로 들어갈 수 있었다. 사진 속 인물은 학회 첫날 아침의 도원과 성준.

AI-Assisted Programming: Applications, User Experiences, and Neuro-Symbolic Techniques

학회 첫 시간 시상 및 안내사항 전달 후 바로 Sumit Gulwani 의 keynote 발표가 있었다. 발표는 프로그램 합성 기술에 관한 내용으로 가득했는데, 프로그램 합성에 있어 예제기반 프로그래밍, 프로그램 수정문제에서의 합성기술 활용, 신경망을 활용하는 등의

접근방법이 해결해 줄 문제들을 소개해주었다. 듣던 중 갑자기 내 눈길을 끌었던 내용이 있는데, OOPSLA'21 Semantic Programming by Example with Pre-trained Models 에서 〈 "India": "New Delhi" = "France": ? 〉와 같은 의미적 대답을 얻기 위해 atomic expression 에 GPT-3 언어 모델을 호출하는 함수를 추가했다는 이야기였다.



인도와 뉴델리의 관계가 프랑스와 파리의 관계라는 의미는 프로그램이나 데이터 형태로서 정해진 것이 아니기 때문에, 저런 정보들이 있다는 사실 자체는 외부 지식 저장소로부터 가져와야 했고, 오늘날 외부 텍스트 자료를 가장 편리하게 저장해둔 저장소는 수천대의 GPU 를 이용해 학습시킨 언어 데이터 모델일 것이다. 저 표현식을 추가함으로써 합성기의 표현력을 높일 수 있을 것이고, 바로 그 점이 내게 인상깊게 다가왔다.
이와는 별개로, 프로그램 합성 관련 발표를 들을 때면 항상 드는 의문이 있다. "그래서 합성 기술이 개발자와 동등한 지위에 오르려면 어떻게 구성되어야 할까" 는 질문인데, 이번에도 속 시원한 답변은 접하지 못했다.

Testing of Autonomous Driving Systems: Where Are We and Where Should We Go?

자율주행 시스템(ADS)에 어떤 테스트가 이뤄지고 있는지 실제 이뤄지고 있는 테스트 방법들을 수집하여 정리하고 연구를 어떻게 진행하면 좋을지 제시하는 연구이다. 예를 들어 외부에서 주어지는 스펙으로는 신호등이 가로인지 세로인지 등의 교통법과 규제사항들이 있을 것이고, 테스트해야 할 소프트웨어는 이미지 센서와 처리만이 아닌 전체 자율주행 시스템으로 규모가 크다고 한다. 하지만 테스트 실무자들은 테스트를 완벽히 수행하기엔 부족한 점이 많다고 했는데, 차가 물에 빠지는 등의 예상치 못한

주행상황들이 정의가 덜 되었고, 테스트에 필요한 시간은 수 주일밖에 없으며, 복잡한 주행상황을 구성하는 작업과 데이터 라벨링이 어렵다는 답변을 받았다고 한다.
발표자가 조사를 통해 얻어낸 많은 사실들을 쏟아내듯 급하게 털어놓아 듣기 어려웠지만, 자율주행 자동차 테스트에 대해 처음이기도 하고, 현업에서 테스팅의 어려움에 대해 정리해 발표해주어 흥미진진하게 들었었다.

Scenario-Based Test Reduction and Prioritization for Multi-Module Autonomous Driving Systems

위의 내용과 이어진, 자율주행 시스템의 테스트 시간을 줄이기 위한 연구이다. 예를 들어 평범한 직선운전을 할 때와 같은 상황에서는 거의 중복되는 내용의 주행기록 정보가 남는다는 관찰에서, 연속된 시간대에서 중복되는 주행기록 정보는 한번만 테스트하게끔 하여 테스트 시간을 줄이려는 접근방법을 제시했다. 이 때 두 인접한 두 프레임의 주행기록 정보가 중복에 가까운지 각 프레임의 feature vector 를 뽑아내어 확인했다고 한다. 앞서 들었던 발표와 이 발표가 같은 세션에 있어 이 연구의 의미가 더 명확하게 느껴졌던 것 같다.

The Evolution of Type Annotations in Python: An Empirical Study

파이썬 3.5 부터 등장한 타입 표기에 대해, 9655 개의 파이썬 프로젝트에서 백만개가 넘는 커밋에서의 타입 표기 변경내역을 조사한 연구이다. 이 연구가 distinguished paper 에 선정된 이유중에는 이 굉장히 큰 조사량이 한몫 하지 않았을까 싶다. 발표는 research question 네 개와 이에 대한 조사 결과에 대한 내용이 주욱 이어졌는데, 잘 관리되는 프로젝트일수록 타입표기 증가속도와 코드 증가속도가 비슷하게 이어진다는 조사결과와, 70% 정도의 타입표기들이 한번 쓰이고 이후에 안 지워지고 있다는 결과를 보여줬는데, 나는 오히려 30%나 되는 많은 양의 타입표기들이 한번 쓰인 뒤로 변경되고 있다는 점에 더 주목해야 한다고 생각한다. 얼마나 타입 추론이 마음에 안들거나 잘못되기 쉬우면 한번 적은 타입 표기를 바꿔 적었어야 했을까 말이다.
여담으로, 학회에서는 commit 내역과 관련된 발표들을 Evolution 이란 이름의 세션으로 구성했는데 내겐 이 단어표현이 좀 생소하게 느껴졌다.

Input Invariants

문법 기반 퍼징에 의미적 제약조건을 주려는 연구이다. 연구동기로 들어준 예시가 유명한 OpenSSL 의 heartbleed 취약점이어서 좀 더 흥미롭게 들었던 것 같다. 이 예시에서, heartbleed 취약점을 일으키려면 네트워크 패킷보다 length 값을 크게 주어야 발견할 수 있는데, 기존의 문법기반 퍼징 테스트 방법들은 이런 제약조건을 지키는 입력을 만들지 못한다고 문제를 보였다. 발표자는 이 문제를 해결하기 위해 바이트 길이 등의 여러 제약조건을 입력할 수 있는 언어를 제시하고, 사용자가 언어로 문법과 스펙을 적어주면 퍼징 테스트에 필요한 입력을 만들어주는 도구를 선보였다.
제약조건을 표현할 수 있는 새 언어를 제시한다는 점에서 수많은 의문점들이 떠올랐다. 제약조건을 얼만큼 많이 표현할 수 있는지, 제약조건을 만족하는 테스트 입력을 얼마나

빠르게 만들어줄 수 있는지 등등. 하지만 제시한 문제가 명확했기에 위와 같은 의문점이 남아있어도 충분히 멋진 연구결과가 될 수 있구나 싶었다.

싱가포르

싱가포르는 항상 더운 나라였고,
우리나라와 비슷하게 치안도 좋은지 저녁의 마리나베이에는 더위를 피할 겸 야경을 구경하러 나온 관광객들이 많이 있었다.
사진 뒤에 높이 서 있는 건물들은 대부분 증권사 건물들이었는데, 멋지다는 생각이 들면서도 한편으로는 저 많은 사람들이 정말 야근중이라 불을 킨 걸까 의문이 들었다.

출국 전 국립미술관에 방문하려다 버스 내릴 곳을 놓쳐 예정에 없이 버스 하차장소와 가까운 국립박물관에 입장했는데 결과적으로는 오길 잘했다 생각한다. 20 세기에 영국의 영향을 많이 받은 채로 비교적 평화롭게 지낸 나라의 지난 100 년간의 문화를 알 수 있었다는 점에서 지금껏 우리나라 역사에 빗대어 보아서는 상상할 수 없는 연대기를 감상할 수 있었다.
사진은 국립박물관 앞에 있는 커다란 고무나무이다.

마치며

이번 FSE 참석으로 연구실 선후배들과 함께 소프트웨어 공학 커뮤니티의 테스팅과 분석 중심으로 여러 연구발표들을 접하고, 싱가포르에 대해 접할 수 있는 기회를 가져 즐거웠다. 학회 참석 기회를 주신 오학주 교수님께 감사드린다.