SIGPL 2024 겨울학교 리포트

Dohyung Kim

February 5, 2024

1 첫째 날(1/31)

몇 년 만에 찾아온 대전과 KAIST는 생각만큼 다르지 않은 모습이였다. 성심당이 유명한 것도 똑같았고, SIGPL 회장에 오자마자 잔뜩 있던 성심당 빵도 똑같이 맛있었다. 다시 튀김소보로를 먹으면 별 감흥이 없을 거라고 생각했는데, 착각이었다. 그렇게 튀김소보로를 먹으며 프로그램 자동 수정에 대한 초청 강연을 들었다. 현재 도운이와 C++ 프로그램 자동 수정을 위한 프로그램 합성 연구를 하고 있는 게 있어 관심 있었던 주제인데, 어떻게 여러 오류들을 유형별로 나눠 각기 처리할지에 대한 이야기를 들을 수 있었다. 아무래도 딥러닝 기반의 연구다 보니 내부에서 패치를 위한 템플릿이 어떻게 생성되는지에 대한 유용한 정보를 얻을 수는 없어서 아쉬웠지만, 우리 방식대로 프로그램 합성을 잘 해낼 수 있다면 이를 어떻게 적용시킬지에 대한 큰 그림을 참고할 수 있었다.

이후 번개 발표에서 흥미로웠던 몇 가지 주제를 정리해보면, Parsing Expression Grammar에서의 Pruning, LLM 코드의 문법 검증, 컴파일러 최적화 검산을 위한 지향성 퍼징, 파이썬 콜 그래프 생성, PAC 모델을 통한 패치 신뢰도 측정, C/Rust 번역 시 출력 결과 보존 등이 있었다. 이후 포스터 발표 시간에 맥주를 마시며 흥미로웠던 주제들에 대한 설명을 들었는데, 기대한 것에 비해서는 실제 결과물이나 내용이 아직 그리 내세우기는 애매한 것들이 많기는 해서 아쉽긴 했다.

별개로 포스터 발표 시간에 이렇게 맥주를 마신 건 처음이라 신기하기도 했는데, 아마 이번 SIGPL에서 학생들 간의 교류를 위해 술을 먼저 좀 마시게 한 게 아닌가 생각된다. 실제로 이후 저녁 식사를 조별로 이동해서 각자 서로 다른 식당에서 하게 되었는데, 이야기를 하다가 비슷한 논조의 말이 꽤나 나왔다. 내가 속한 11조는 아카바의 식탁이라고 하는 퓨전 양식당에서 저녁을 먹었으며, 대학원생 생활에 대한 이야기를 나누다가 헤어졌다.

2 둘째 날(2/1)

아침에 KAIST로 가는 길에서 호텔에서 나와 걷고 있자니, 아무래도 오르막길이 없는 등굣길은 참 편하다는 걸 새삼 느낄 수 있었다. 월평공원에서 감전을 따라 도는 산책로를 따라 걷다가 SIGPL 회장에 도착했는데 아쉽게도 둘째날에는 튀김소보로를 주지 않았다. 첫 번째 초청강연은 임베디드 시스템에서 소형 메모리 기반의 칩에 어떻게 ML용 컴파일러를 실어야 하는지에 대한 것이었는데, 어떻게 중복되는 것들을 최대한 배제하며 생성되는 코드와 계산을

쪼개가며 처리하고 다시 넣고 할 수 있는지를 구체적인 코드나 하드웨어 구조를 통해 설명했다. 좀 더 세부적으로는, 각각의 프로그램이 언제 메모리를 사용해야 하나를 분석할 필요가 있으며, 어떤 과정이 주로 반복되며 어떤 것을 간단히 대체할 수 있는지를 분석해 이 분석한 내용들을 바탕으로 컴파일 과정에 동적으로 메모리를 할당하며 최적화를 진행한다.

또한 다음 초청 강연에서 근사 컴퓨팅을 통해 비교적 덜 정확한 계산을 통해 성능을 더 빠르게 개선하는 컴파일러 기법을 설명하는 것을 흥미롭게 들었다. 최근 정적 분석에서도 꼭 안전성(soundness)만을 보전하는 것이 아니라 디버깅 활용을 위해 빠르게 오류를 잡을 수 있도록 정확도를 타협하는 정적 분석기를 제안하는 걸 읽어서 그렇기도 하다. 이 때 중요한 것은 어떻게 근사 시킨 계산이 기존의 완전한 계산과 비교해 정확도를 유지하는지를 측정하는 걸 테다. 강연에서 인상적이었던 건 어떻게 동일한 동작을 하는 두 상위 프로그래밍 언어 프로그램이 근사 시킨 계산에서는 전혀 다른 동작을 하는지를 파악하는 문제였다. 하위 언어를 활용하도록 곧장 진입하는 것보다는, 상위 언어에서 최대한 해결할 수 있도록 노력하되 극복하기 힘든 문제를 직면했을 때에만 단계적으로 깊게 들어가는 실용적인 전략이 필요하다.

다른 강연에서는 크게 감흥을 받지 못했지만, 빅 데이터 스트리밍 엔진의 성능 평가에 대한 강연이 다른 의미로 상당히 인상적이었다. 전자 신호가 어떤 식으로 전달되며 이 신호의 서로 다른 시간대에서의 지연도를 측정하기 위해 여러 편미분 방정식을 풀던 걸 보며, 학회 주제와 얼마나 맞는지는 잘 모르겠지만 이러한 계산을 바탕으로 정적 분석을 하는 것이 상당히 인상적이었다. 이후 저녁을 먹은 뒤 뱅킷에서 나온 맥주를 마시며 랩 전원이 함께 모여 술을 마시며 일종의 뒷풀이 시간을 가졌다.

3 셋째 날(2/2)

마지막 날에는 들고 갈 수 있게끔 성심당 과자들이 담긴 도시락을 받았다. 아쉽게도 튀김소보로는 딱히 들어 있지 않았다. 마지막 초청 강연은 웹 코드의 코드 속성 그래프를 탐색하는 쿼리를 통해 코드의 취약점을 빠르게 파악하는 방식에 대한 것이었는데, 기본적으로 취약점이 발생할 수 있는 특정 함수들을 찾아내고 이 함수에 들어가는 정보의 흐름을 파악해 위험한 입력이 들어가는 것을 방지하는 정적 분석이다. 이 기존의 방식만으로는 큰 규모에서 제대로 취약점을 탐색할 수 없어, 강연에서 소개한 새로운 방식은 특이하게도 취약점이 발견되는 그래프 쿼리의 부분 그래프 형태들을 수집해, 탐색하고자 하는 큰 PHP 코드 속성 그래프에서 이런 부분 그래프들을 매칭하는 것으로 빠르게 취약점이 있을 것으로 예상되는 부분들을 찾아내는 것이다. 상당히 특이한 접근법과, 이를 실제로 적용하는 과정에서 발생하는 문제점들을 듣는 것도 꽤나 도움이 됐다.

SIGPL이 끝난 후 단체로 남은 시간 동안 대전의 명물을 찾아 돌아다녔다. 점심으로는 칼국수와 두부 두루치기를 먹은 후, 성심당에 몰려가 빵을 사왔다. 역시나 성심당 본점에는 사람이 가득했는데, 신기하게도 그렇게나 많은 사람들이 줄을 잔뜩 서고 있는데도 불구하고 빵집이다보니 사람들이 빠르게 빠져 얼마 기다리지 않은 채 빵을 사서 나올 수 있었다. 성심당에서 유명한 튀김소보루와 부추빵도 맛있었지만, 대전 사람들에게 거의 식사용으로 사가는 빵에 가깝다는 이야기를 들어서 사온 다른 몇 가지 빵들도 전부 맛있었다. 오랜 명성에는 이유가 있다.

이후 돌아가는 길에 SIGPL에서의 경험을 정리해보았는데, 상당히 흥미로운 주제에 저마다 다른 방식으로 접근하며 LLM을 신기한 방식으로 도입해보는 이들도 많아 여러 가지로 자극되었다. 금년도 겨울 SIGPL은 더욱이 강연 뿐 아니라 여기에 참여하는 학생들의 연구 및 소통이 꽤나 중요시된 자리다보니 신기하기도 했다. 사실 같은 연구실이면 몰라도, 다른 연구실의 학생들과 소통하거나 연구를 같이 할 일이 그리 많지는 않다. 같은 학교에서도 별로 이야기하지 않는 경우가 허다하다. 다만 이번 SIGPL에서는 이런 문제를 조금 해소하는 겸 잘 어울릴 수 있도록 식사를 비롯한 여러 활동을 진행하고, 포스트 세션이나 뱅킷에서 맥주를 제공하는 등 최근 SIGPL과는 꽤나 다른 모습을 보여주었다. 덕분에 다른 학생들이 대학원에서 어떻게 생활하고 있는지를 좀 더 진지하고 다양하게 들어볼 수도 있어서 학업적 측면과는 또 다른 의미에서 흥미로웠던 시간이었다.