

FSE 2025

» TRONDHEIM, NORWAY



HYUNJOON KIM

JUNE 2025

INTRODUCTION

FSE는 소프트웨어 공학 분야에서 ICSE와 더불어 최고 수준의 학회로, 매년 수많은 연구자들이 본인의 연구 성과를 발표하고 새로운 아이디어를 얻기 위해 이 학회를 찾는다. 나는 이번에 FSE'25 Demonstration track에 민석이와 경호 형과 함께 작성한 JSSpecVis 논문이 역선택되어 연구실의 지원을 받아 참석할 수 있었다. 이번 논문의 1저자인 민석이가 발표를 담당하여 나는 비교적 편하게 학회를 즐길 수 있었는데 참석의 기회를 주신 박지혁 교수님께 감사드리며, 이 글을 통해 학회에서 경험하고 느낀 내용들이 잘 전달되고 도움이 되길 바란다.

이번 FSE'25는 ISSTA'25와 같이 노르웨이 트론헤임에서 23일부터 27일까지 총 5일에 걸쳐 진행되었다. 나와 민석이는 FSE가 열리는 월요일부터 수요일까지 3일 동안 참석하였고, 학회 전후로는 노르웨이의 문화와 자연을 경험할 기회를 가질 수 있었다.

일주일 전 서울에서 개최된 PLDI'25 참석으로 국제 학회의 분위기에 어느 정도 익숙해졌다고 생각했지만, 첫 해외 학회 참석이라는 사실에 설렘과 긴장이 공존했다. 환승을 제외하고도 16시간에 달하는 긴 비행 끝에 도착한 노르웨이 트론헤임은 도시의 풍경이 정말 아름다웠다. 노르웨이의 물가는 예상보다 훨씬 높아 적응이 필요했지만, 학회장 주변의 아름다운 북유럽 도시 풍경과 친절한 현지인들 덕분에 즐겁게 학회 전후를 보낼 수 있었다.



Figure 1: 트론헤임 시내와 니다로스 대성당

학회 첫날 등록을 마치고 둘러본 회장은 PLDI보다 훨씬 큰 규모였다. 소프트웨어 공학이 프로그래밍 언어보다 더 큰 분야임을 실감할 수 있었고, 그만큼 다루는 연구 주제들도 매우 다양했다. LLM을 활용한 연구들이 지나치게 많은 비중을 차지하고 있어 아쉬움이 남았지만, 퍼징과 테스팅 관련 발표들도 충분히 많아 재밌게 들을 수 있었다. 이어지는 섹션에서는 3일간의 학회 기간 동안 특히 인상 깊었던 퍼징과 테스팅 관련 발표들 몇 개를 소개하고자 한다.

INTERESTING TALKS

MENDELFUZZ: THE RETURN OF THE DETERMINISTIC STAGE

퍼징 분야에서 오랫동안 당연하게 여겨왔던 선택에 도전한 이 발표는 이번 학회에서 들은 퍼징 발표 중 가장 실용적이면서 인상 깊은 발표였다. 대부분의 현대 퍼저들이 성능 문제로 기본 설정에서 비활성화했던 deterministic stage를 다시 살펴보고, 이를 효율적으로 되살려낸 연구였다.

AFL과 같은 퍼저에서 채택하는 mutation 전략은 크게 deterministic stage와 havoc stage로 나뉜다. 전자는 각 바이트를 체계적으로 하나씩 하나씩 변형하는 반면, 후자는 여러 mutator를 동시에 적용한다. 현재 대부분의 상용 퍼저들은 성능상의 이유로 havoc stage를 기본으로 채택하고 있다. Deterministic stage가 전체 실행 시간의 대부분을 차지할 정도로 비효율적이기 때문이다.

그러나 저자는 deterministic stage의 비효율성을 단순히 받아들이는 대신, 그 원인을 깊이 있게 분석했다. 핵심 발견은 놀라웠다. 단지 0.5%의 바이트가 전체 경로 발견의 84%에 기여하고, 20%의 시드가 coverage 증가의 83%를 담당한다는 것이었다. 즉, deterministic stage의 대부분의 작업이 실제로는 불필요했던 것이다.

이러한 관찰에서 이 연구는 execution trace를 분석하여 중요한 바이트와 시드만 골라내고, 이들에만 deterministic stage를 적용하였다. 결과적으로 시간은 10% 이하로 줄이면서 효과는 유지할 수 있었다. 또한 실제로 기존 deterministic 퍼저들보다 10% 더 많은 coverage를 발견하고, 62.33% 더 많은 버그를 찾았다.

발표를 들으며 느낀 건, 비효율적이라고 판단되어 외면받던 기술도, 그 비효율의 원인을 정확히 파악하고 선택적으로 적용한다면 오히려 더 나은 결과를 낼 수 있다는 것이다. 이 연구가 이제 AFL++의 기본 모드로 채택된 것도 이런 접근의 실용적 가치를 인정받은 결과라고 할 수 있다.

UNITCON: SYNTHESIZING TARGETED UNIT TESTS FOR JAVA RUNTIME EXCEPTIONS

이 연구는 특정 위치에서 런타임 예외를 트리거하는 단위 테스트를 자동으로 합성하는 연구로 내가 최근에 관심을 갖고 있던 문제와 유사한 연구였다. 문제는 명확하다: 기존 EvoSuite나 Randoop 같은 테스트 생성 도구들은 전체 프로그램의 coverage를 높이는 데 집중하지만, 특정 위치의 버그를 재현하는 예제를 만드는 데에는 효과적이지 못하다는 것이다.

UnitCon의 핵심은 정적 분석을 전략적으로 활용해 탐색 공간을 pruning하고 우선순위를 지정하는 것이었다. 사실 이러한 방법론 자체는 program synthesis 분야에서 흔히 보는 접근이다. 동일한 행동을 하는 중복된 후보들을 제거하고, 타겟 에러를 트리거할 가능성이 높은 후보에 우선순위를 부여하는 방식.

다만, 이러한 분석을 scalable하게 해결한 것이 인상적이었다. 정적 분석을 그대로 적용하면 매우 느릴 텐데, UnitCon은 SMT solver에 전달할 중요한 정보를 최소한으로 추출하여 이를 해결했다. 예를 들어 JSqlParser의 경우, 70개의 메서드 중에서 실제로 NPE와 관련된 소수의 메서드만 식별하고, null과 비교하는 등의 핵심 조건만을 추적한다. 이렇게 데이터에 대한 지식을 활용해 꼭 필요한 정보만을 선별하는 것이 통찰력이 빛나는 지점이라고 생각했다.

실험 결과도 매우 설득력 있었다. 198개의 벤치마크에서 104개의 버그를 재현했고, 이는 기존 도구들보다 1.2-3.6배 많은 수치다. 더 나아가 실제 오픈소스 프로젝트에서 21개의 새로운 NPE 버그를 발견했다는 점은 이 연구의 실용성을 잘 보여준다.

이 발표를 들으며 든 생각은 좋은 도구는 이론과 실무의 균형을 잘 맞춘다는 것이었다. 복잡한 synthesis 알고리즘을 그대로 적용하는 것이 아니라, 테스트 생성이라는 특정 도메인의 특성을 이해하고 필요한 부분만 효율적으로 추출하는 방식. 데이터를 열심히 케이스 스터디해야겠다는 생각이 다시금 들었다.

DE-DUPLICATING SILENT COMPILER BUGS VIA DEEP SEMANTIC REPRESENTATION

개인적으로 이 발표를 들으며 복잡한 감정이 교차했다. 논문의 내용이 새롭다기보다는 이전에 비슷하게 버그를 일으키는 코드들을 deduplication 하는 연구를 진행하다가 중단한 적이 있었기 때문이다.

저자들이 이 아이디어를 최초로 제시한다고 말할 때, 과거에 충분한 시간과 노력을 투자하지 못했던 나 자신이 떠올랐다. 하지만 한편으로는 spec 기반의 deduplication이라는 방향성은 여전히 유효하다는 생각도 들었다. 이 연구가 learning 기반의 접근에 중점을 둔다면, spec 기반 접근은 또 다른 차별화 지점이 될 수 있을 것이다.

아이디어는 여러 사람이 독립적으로 생각할 수 있지만, 그것을 끝까지 밀고 나가 완성도 있는 연구로 만드는 것이 진짜 실력이라는 교훈을 얻었다. 좋은 동기부여가 되는 발표였다. 연구의 내용도 재미있으니, 굳이 여기서 다루지는 않겠지만 읽어보기를 추천한다.

BEYOND THE TALKS

학회 일정 외의 시간에는 트론헤임의 아름다운 도시 풍경을 즐기며 산책을 주로 하였다. 특히 백야 현상으로 밤 11시에도 해가 지지 않아 독특한 경험을 할 수 있었다.

노르웨이의 물가는 예상을 훨씬 뛰어넘는 수준이어서, 간단한 식사에도 상당한 비용이 들었다는 점이 인상적이었다. 한 끼에 3-4만원은 기본이었고, 커피 한 잔도 만원에 가까운 가격이었다. 아이러니하게도 학회장에서 제공된 식사는 만족스럽지 못했다. 특히 전주에 참석했던 PLDI의 케이터링과 비교하면 그 차이가 더욱 두드러졌다. 높은 물가에 비해 학회 케이터링의 퀄리티가 아쉬웠던 것은 모두들 공감했다.



Figure 2: 맛있는 밥은 언제나 중요하다

시차 적응 문제로 학회 talk들을 듣고 난 후에는 일찍 휴식을 취해야 했다. 7시간의 시차는 생각보다 크게 체감되었고, 특히 저녁 시간대의 피로감이 상당했다. 오후 세션에서 집중력이 떨어져 아쉬움이 남았지만, 자고 일어나도 여전히 밝아 시간 감각이 무뎌져 오히려 적응에는 도움이 되었던 것 같다.



Figure 3: 민석이의 발표

특별히 기억에 남는 것은 함께 참석한 민석이의 논문 발표였다. 준비 과정부터 실제 발표까지 지켜보면서, 첫 국제 학회 발표임에도 불구하고 침착하게 잘 해내는 모습이 인상적이었다. 발표 전의 긴장감과 발표 후의 안도감, 그리고 질문들에 대한 응답을 무사히 마친 후의 성취감까지 모든 과정을 가까이서 지켜볼 수 있었다. 동료의 성공적인 발표를 지켜보며 내 스스로도 지금 연구를 더욱 열심히 진행해서 이 무대에 서고 싶다는 강한 동기부여가 되었다.

CLOSING THOUGHTS

이번 FSE'25 참석은 여러 면에서 좋은 자극이 되었다. 특히 발표들을 들으며 데이터를 꼼꼼히 분석하는 것의 중요성을 다시 한번 느꼈고, 시작한 아이디어는 끝까지 밀고 나가야 한다는 교훈을 얻었다.

첫 해외 학회라 미숙한 부분도 많았지만, 다음에는 더 적극적으로 참여하고 싶다는 욕심이 생겼다. 무엇보다 민석이의 발표를 보며 나도 다음번에는 직접 발표하는 기회를 만들고 싶다는 동기부여를 받았다.

지원해주신 박지혁 교수님께 감사드리며, 이번 경험을 토대로 현재 진행 중인 연구를 더욱 열심히 진행하겠다는 다짐을 하며 이 글을 마친다.