

Oxford Talks & ICST 2018

Sooyoung Cha

Korea University

2018 April 7-14



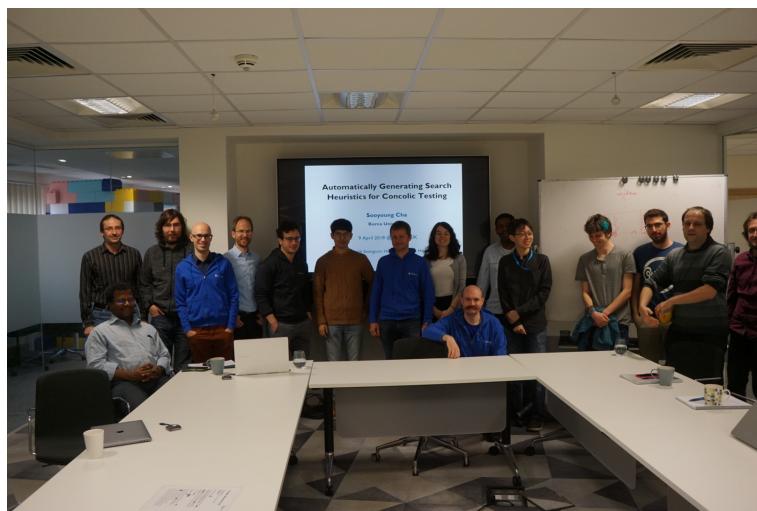
1 개요

저는 4월 9일 옥스포드 대학과 DiffBlue 회사에 연구주제를 발표하러 영국 옥스포드에 다녀 왔습니다. 또한 4월 10일에서 12일까지 스웨덴 Västerås에서 열린 ICST(IEEE Conference on Software Testing, Validation and Verification)에 참석했습니다. 두 번의 발표와 테스팅 학회에서 얻은 경험들이 잘 공유되기를 기대합니다.

2 발표

2.1 DiffBlue Talk

먼저 첫 발표는 Daniel Kroening 교수가 CEO로 있는 Diffblue라는 테스팅 회사에서 시작했다. 발표 자체는 준비를 많이해서 걱정되지 않았지만, 발표 중간 또는 끝에 있을 많은 질문들과 발표의 청자들이 모두 테스팅 전문가라는 사실이 부담되었다. 발표는 총 30분 정도를 준비하였고, 초반 10분 정도는 Concolic Testing의 배경지식을 이야기하는 것이 계획이 있다. 하지만 예제를 설명하려고 하는데 듣는 사람들이 모두 너무 잘 아는 눈으로 나를 쳐다보았다. 그래서, 예제 설명은 빠르게 skip해도 되냐고 물어보고 청자들이 관심있어할 만한 곳으로 빠르게 넘어갔다. 즉 나의 발표 시간은 이제 20분으로 줄어든 것이다. 그리고 나서, Motivation을 설명하는데, 모두가 수긍하고 공감하고 있다는 생각이 들었다. 이번 연구는 개인적으로 굉장히 Motivation이 명확하다고 생각했고, 그것이 좋은 연구의 꼭 필요한 것이라는 것을 다시 한번 느꼈다. 그 다음부터 발표는 일사천리로 쉽게 잘 흘러갔던 것 같다. 마지막 슬라이드의 Tool 이름을 맞추는 것까지 좋은 분위기로 잘 끝났다. 이제는 질문시간이다. 질문은 최소 내 발표시간보다는 오래 받았다.



Key Questions

1. 큰 프로그램 (e.g., window)을 나눠서 각 부분들을 테스팅한다고 할 때, 너의 기술이 유용하지 않을까?

2. 동일한 프로그램에 대해서 탐색전략을 만들때마다 매번 동일한 전략이 만들어지는가?

3. feature들을 디자인하는 과정에서 redundant한 feature들을 어떻게 판단하였는가?

1번 질문은 내가 전혀 생각해보지 못한 방향이라 처음에 잘 알아듣지 못했던 질문이었다. 질문의 요지는 한번에 전체 프로그램을 테스팅하기 어려운 큰 프로그램 (예:window)에 대해서, 각 부분들을 나눠서 테스팅할텐데, 그 때 각 부분을 최적화된 탐색전략을 생성해서 테스팅하면 좋지 않을까라는 질문이었던 것이다. 우리 기술이 실제 산업에서 유용하게 쓰일 수 있는 좋은 시나리오라고 생각한다. 2번 질문은 개인적으로 아주 날카로운 질문이라고 생각한다. 우리 parameter search 알고리즘에는 randomness가 존재하기 때문에 경향은 있지만, 매번 동일한 전략(동일한 parameter value)이 생성되진 않는다. 우리 기술의 약점으로도 볼 수 있는 것이기 때문에, 잘 보완해야되는 점이라고 생각한다. 40개의 feature에 대해서도 많은 질문들이 있었다. 예를 들어, 질문3은 “처음엔 40개보다 많은 feature들이 있었는데 그 중에서 redundant feature들을 제거했다”라고 내가 발표한 것에 대해서 질문한 것이다. 사실 구체적인 기준을 정하고 한것은 아니고, “직관 + feature marking이 너무 특정 branch들에 대해서만 될 때에 제거했다”라고 대답했다. 그 외에도 내가 예상하지 못한 여러 질문들이 나와서 긴장도 했었고 신이 나기도 했다. 반대로 내가 예상한 질문들이 나왔을 때 반가운 나머지 난 “good question”이라는 말을 덧붙였던 것 같다. Diffblue 회사에 대한 나의 인상은 ‘이 사람들은 매우 똑똑하고 내 연구를 거의 다 이해한 후, 질문들을 하는구나’였다.

2.2 Oxford Talk

두번째 발표는 옥스포드 대학교 컴퓨터학과에 있는 대학원생들 또는 박사들을 대상으로 진행됐다. 참석인원은 약 30명 정도였던 것 같다. 나를 친절하게 맞아준 Ohad Kammar 박사님과 2시간 전에 발표를 한번했던 경험 덕분에 더 여유있고 차근차근 설명할 수 있었던 것 같다. 내 세미나의 분류가 Programming Language로 되어 있어서 PL 연구실 사람들이 많이 왔고 모두 Concolic Testing에 대해서 어느정도 지식이 있는 상태인 것을 확인했다. 이 발표에서도 재미있는 질문들이 많이 나왔고 꽤 활발하게 토론을 진행했던 기억이다.



Key Questions

1. Transfer Learning을 시도해보았는가?
2. Optimization을 다른 쪽(e.g., bug-finding)으로 시도해본 적이 있는가?
3. Probabilistic Programming의 Monte Carlo Method와 너의 방법과 비슷하다고 생각하지 않는가?

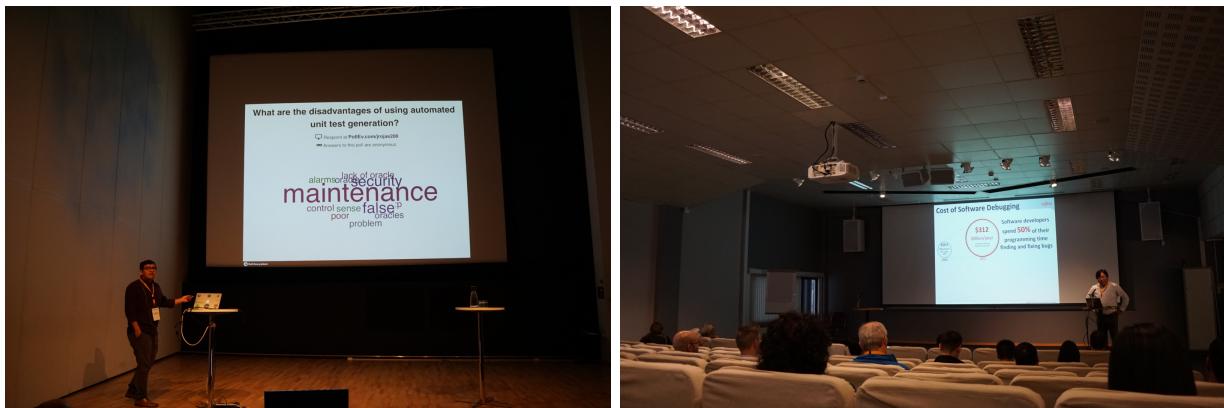
질문 1,2번은 모두 ICSE 논문을 저널로 확장할 때, 좋은 대상이라고 생각한다. 20분 발표를 듣고, 바로 나와 같은 생각을 한다는 것이 놀라웠다. 3번 질문은 Probabilistic Programming을 하는 여러 학생들의 공통 질문이었다. Ohad Kammar 박사님도 동일하게 물어보려고 했던 질문이라고 했다. 발표가 끝난 후에는, 시간이 되는 사람들끼리 한식당에 가서 같이 밥먹으면서 이런 저런 연구 이야기들을 하였다.

두 발표에서 한가지 얻은 팁은 발표 내의 적절한 퀴즈나 웃음 포인트는 발표를 더 집중되고 편안하게 들을 수 있게 해준다는 것이었다. 예를 들어, 졸업생 그림은 육스포드 대학원생들에게 큰 웃음과 공감을 주었다. 그리고 다른 하나의 팁은 ‘Motivation을 설명할 때 신경을 더 쓰자’입니다. 연구의 Motivation을 청자들에게 납득시킬 수 있는지 여부가 발표의 성공을 결정한다고 생각이 들었다.

3 ICST 2018

3.1 학회 참석

ICST는 테스팅에 전문화된 학회여서 나에게 꽤 유익할 것이라는 기대를 품고 참석하였다. 발표 들을 들으면서 느낀 첫 느낌은 테스팅의 분야가 매우 다양하다는 점이었다. 예를 들어, Visual Testing은 나에겐 생소한 분야였고, 이 세션에서 류석영 교수님 연구실 학생이 HTML5에 있는 Visibility Faults를 탐지하는 기술을 발표하기도 했다. 그래서 기대와 달리 Concolic Testing을 이용한 연구들은 그렇게 많지 않았다.



ICST Research Paper Talks

발표 발표시간은 한 발표당 약 25분정도로 꽤 충분히 주어졌다. 한가지 흥미로운 발표 방식은 발표 도중에 저자가 질문을 던지고, 실시간으로 청자들이 핸드폰으로 그 질문에 대답한 내용들이 발표 슬라이드에 반영되는 거였다. 나를 비롯한 청자들의 집중을 이끌 수 있었고, 시간이 충분히 있을 때는 시도해볼만한 발표방식이라고 느꼈다. 한국에서는 총 두 개의 발표를 했다. 특히 홍신 교수님이 발표하신 연구는 Concolic Testing 기반이고 도구는 CREST를 이용했기 때문에, 더 즐겁고 쉽게 이해할 수 있었고 질문도 할 수 있었다. 또한 내가 진행해보고 싶던 주제인 “강화학습 + 테스팅”으로도 연구가 발표되어서 흥미롭게 들었다.

Banquet Banquet 장소는 학회 장소에서 차로 40분정도 떨어진 Munktell 박물관이었다. 스웨덴의 1900년대 초반에 이용되던 트렉터들을 볼 수 있는 박물관이었다. 1시간 정도 구경을 한 후, 트렉터들에 둘러쌓인 박물관 안에서 저녁을 먹는 색다른 경험을 할 수 있었다. 또한 각 테이블별로 주어진 퀴즈들을 푼다거나, 중간중간 PC Chair들이 진행하는 만담 등은 Banquet을 매우 즐겁게 보낼 수 있게 해주었다.



Munktell Museum

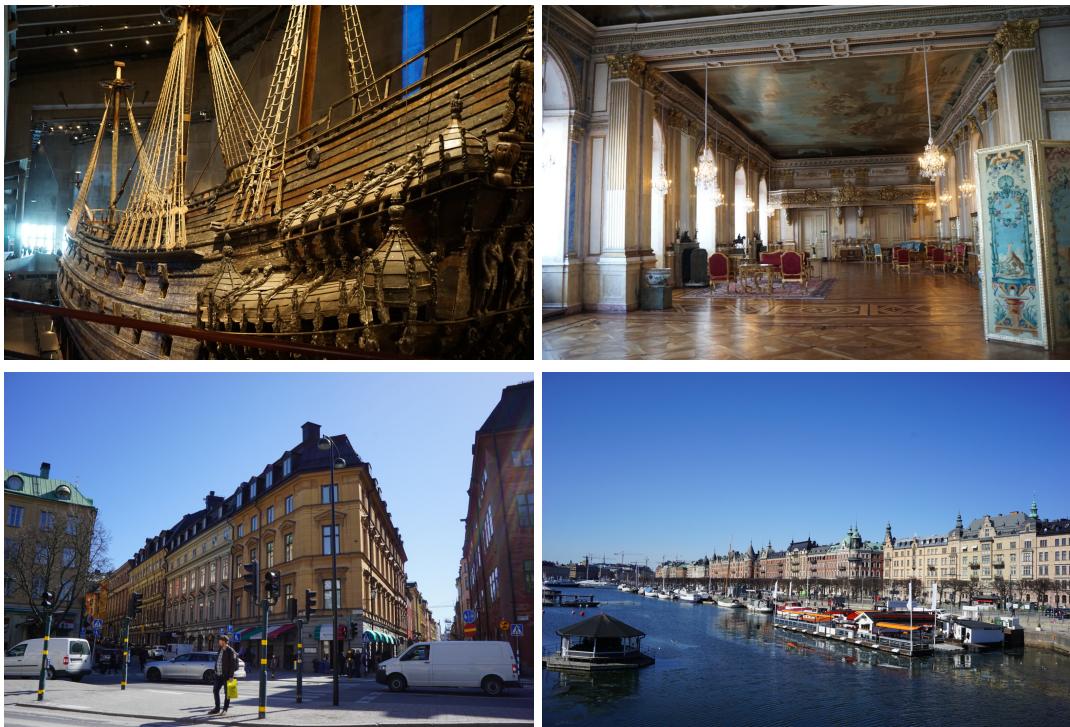
3.2 논문 소개

QBE: QLearning-Based Exploration of Android Applications 이 논문은 안드로이드 GUI 테스팅을 효과적으로 하기 위해 강화학습을 적용한 첫 연구이다. 제안한 테스팅 도구 QBE는 각 어플리케이션의 여러 GUI action들을 시도해보는 테스팅을 수행할 때, 랜덤하게 GUI action을 고르는 것이 아니라 objective에 최적화시킨 GUI action들의 확률 분포를 이용해서 action을 선택한다. 예를 들어, objective는 activity coverage 또는 오류 검출율을 최대화 시키는 것이 될 수 있다. 그리고 GUI action들의 확률 분포는 잘 알려진 강화학습 알고리즘인 Q-learning 을 이용해서 구한다. 실험적으로 기존의 최신 GUI 테스팅 도구들(A^3E , Dynodroid, PUMA, SwiftHand, and Sapienz)과 비교를 통해, 높은 activity coverage와 오류 검출율을 달성한 것을 보여주었다. 신선향을 주는 방법론(강화학습 + 테스팅)이 눈에 띄었던 논문이었다.

Invasive Software Testing: Mutating Target Programs to Diversify Test Exploration for High Test Coverage 이 논문은 distinguished paper로 선정된 논문으로 홍신 교수님이 발표하셨다. 이 논문의 motivation은 원본 프로그램의 많은 변종(variants)들을 테스팅하면서 얻은 귀중한 정보를 이용하지 못하는 것이다. 이를 위해, 테스팅을 통해 얻은 유용한 정보들을 추출하고, 그 정보로부터 ‘guidepost’라는 것을 생성해낸다. 마지막으로 이 ‘guidepost’를 만족하는 실행을 하도록 가이드 함으로써, 높은 test coverage를 성취하려는 것을 목표로 한다. 실험적으로 테스팅 도구 ’DEMINER’는 3개의 벤치마크(Busybox-ls, Busybox-printf, GNU-find)에 대해서 conventional concolic testing보다 24.7% 높은 branch coverage를 달성할 수 있었다.

4 스톡홀름

학회가 끝나고 하루정도 스톡홀름을 돌아다닐 시간이 있었는데 매력이 많은 도시였다. 스톡홀름 중앙역 근처에 모두 볼 것들이 모여있어서 걸어다니면서 도시를 느낄 수 있었다. 실제로 한 8시간은 걸은 것 같다. 처음 사진은 vasa 박물관으로 17세기부터 무려 333년동안 바다 속에 가



라앉아있던 배 한척이 전시되어 있다. 규모는 사진에 담기지 못할정도로 엄청나게 큰 배였고, 인상깊었다. 나머지 사진들은 각각 스톡홀름 궁전, ‘감라스탄’이라고 불리는 구 시가지, 그리고 그냥 이쁜 풍경이다. 스톡홀름은 저녁 8시가 넘어도 해가 떠 있어서 돌아다니기 좋았고 날씨가 너무 화창해서 좋았다. 다시 한번 가보고 싶은 도시이다. 곧 간다 또.

5 마무리

이번 두 번의 발표는 내가 더 성장할 수 있었던 잊지 못할 아주 귀중한 경험이 되었다. 부담감은 좀 있지만 발표를 할 때마다 느끼는 것은 내 연구를 다른 사람들에게 전달하는 일은 늘 재미있다는 것이다. 앞으로도 좋은 기회를 많이 만들도록 더 노력해야겠다. 끝으로, 제가 더 성장할 수 있도록 좋은 기회를 주신 오학주 교수님께 정말 감사드립니다. 다음엔 더 좋은 연구를 통해 직접 초청받아서 발표하러 다녀오겠습니다!

