
PLDI 2016 여행기

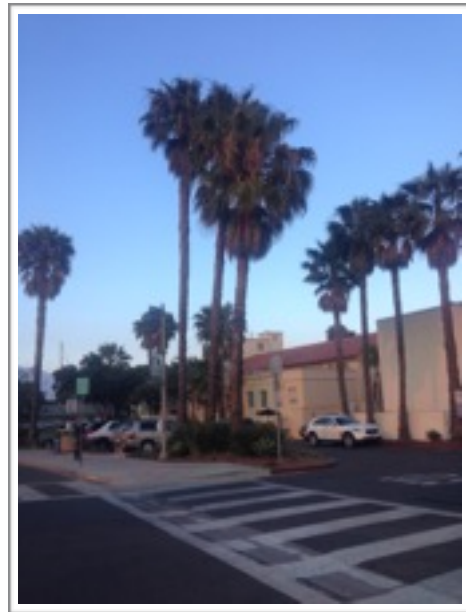
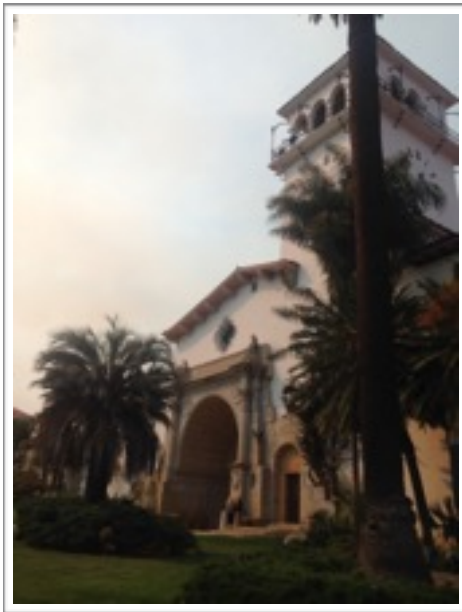
미국, 산타바바라

2016년 6월 12일 - 2016년 6월 18일
서울대학교 허기홍



들어가며

POPL 과 함께 그 해 프로그래밍 언어 분야에서 가장 우수한 논문이 채택되는 학회인 PLDI. 2012년과 2014년에 이어 이번에 세 번째로 참가하였다. 소프트웨어 무결점 연구센터를 통해 알게된 한국 연구자들을 비롯하여 과거 PLDI 나 여타 학회에서 만난적이 있는 사람과도 반갑게 인사를 나누고 교류할 수 있는 좋은 기회였다. 특히 이번에는 PLDI 학생 연구 경진 대회에 참가하여 포스터 발표도 하는 기회를 얻어서 즐거움을 더했다. 한국에서 열심히 하던 연구를 잠시 접어 놓고, 여러 나라에서 온 훌륭한 연구자들에게 새로운 자극을 받은 좋은 경험이었다. 조용한 바닷가를 거닐며 많은 생각을 할 기회를 얻은 것도 값진 선물이었다.



포스터 발표

Efficient Global Octagon Analysis in the Big Data Era

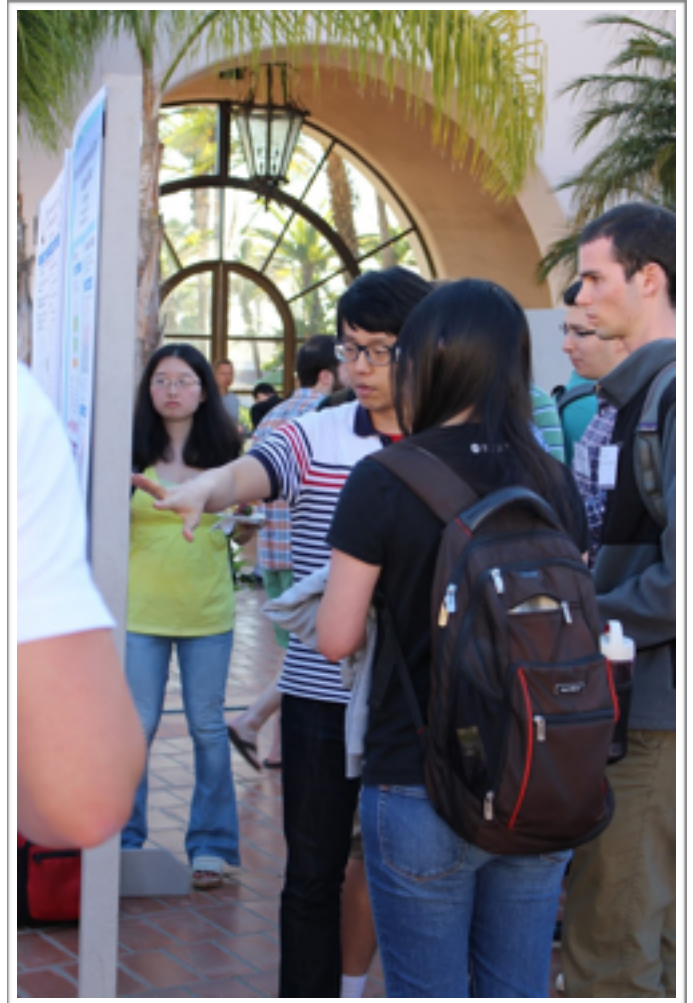
허기홍, 오학주, 양홍석, 이광근

이번 PLDI 에 학생 연구 경연 대회에 참가하여 포스터 발표를 하러 PLDI 로 향했다. 캘리포니아 산타바바라 해변이 보이는 호텔 앞 지는 석양 아래에서, 환영식과 겸해 열린 포스터 발표회였다. 이 연구는 몇 달 전 SAS 학회에 “Learning a Variable-Clustering Strategy for Octagon from Labeled Data Generated by a Static Analysis” 이라는 제목으로 제출한 것인데, 다행히도 출국 전날에 채택되었다고 연락이 왔다. 덕분에 더 자신감을 갖고 즐겁게 발표를 할 수 있었다.

이 연구에서는 기계학습을 이용하여 팔각 관계 분석 (octagon relational analysis) 를 꼭 필요한 부분에만 적용하는 방법을 제시하였다. 관계 분석은 비관계 분석에 비해 정확하지만 비용이 너무 커서 곧바로 적용하기가 어렵다. 따라서 필요한 부분만 잘 선별해서 적용하는 것이 중요한데, 관계 분석이 꼭 필요한

부분만 선별해내는 방법은 2년전 PLDI 에 제출한 바 있다. 그 때는 같은 목표를 달성하기 위해서 정적분석을 사용하였다. 관계 분석의 효과를 쉽고 빠르게 가늠하는 예비분석을 통해서였다. 예비분석을 이용하면 관계 분석을 효율적으로 사용할 수 있지만, 예비분석 자체의 비용도 여전히 큰 것이 문제였다. 그래서 이 연구에서는 예비분석을 학습과정에서 선생님이 사용하고 기계학습 알고리즘을 이용해 관계 분석이 필요한 부분만 배우는 방법을 제시한 것이다. 이를 통하면 기존 방식보다 수십배 빠르게 예측할 수 있어 효율적이다.

포스터에 와서 이야기 나눈 여러 사람들이 기계학습으로 정적분석 (예비분석)을 배운 것에 흥미를 보였다. 연구를 하면서 나도 그랬다. 기계학습을 통한 방식의 결과가 좋아질수록 재미가 있었지만 묘한 기분도 들었다. 우리가 열심히 공부한 정적 분석, 요약 해석 기술이 쓸모 없어지는 것이 아닐까? 알파고가 이세돌을 이긴 충격파가 우리 분야에까지 밀려온 느낌이었다. 그런데 알파고의 승리 이후로 오히려 다시 바둑 열기가 고조되고 기원이 북적거렸듯이 우리 분야도 그러리라 확신한다. 사람이 탄탄한 이론을 통해 잘 디자인한 정적 분석을 기계가 데이터를 이용해 배울 것이다. 그러면 그동안은 상상하지 못했거나 어려워 했던 많은 문제를 쉽게 푸는 새로운 기회가 열릴것이라 기대한다.



초청 강연

Probabilistic Learning from Big Code

Martin Vechev

요즘 기계학습을 이용한 프로그래밍 언어 연구로 각광받고 있는 Martin Vechev 의 초청강연이다. 나와 관심사가 비슷하여 최근에 이 사람의 논문을 읽고 있었는데 이번에 초청강연을 한다고 해서 기대를 품고 들었다. 과거 수년간 POPL, PLDI, ICML 등에 낸 논문에 관한 전반적인 이야기와 경험, 계획등을 이야기해 주었다. 가장 유명한 연구로는 난잡하게 쓰여 있는 자바스크립트 코드를 깔끔하게 바꾸어주는 JSNice 가 있고, 그 이외에도 코드 자동완성, 동시성 프로그램 분석, 그리고 효율적인 팔각 관계 분석 라이브러리 등 다양한 분야를 연구하는 사람이다.

가장 인상깊은 내용은 마지막 즈음에 보여준 한 슬라이드였는데 과거, 현재, 미래의 궤적이 다 들어있었다. 처음에는 PL, SE 분야의 구체적인 문제 (JSNice, 코드 자동완성 등) 를 기계학습으로 푸는 것에서부

터 시작했는데 하다보니 자기들의 입맛에 맞는 기계학습 방법을 구상하게 되었고 결국엔 기계학습 학회에 논문을 낼 정도로 그 분야에서 업적을 이루어 냈다고 한다. 그리고 앞으로는 이렇게 개발한 기계학습 기술을 바탕으로 다시 구체적인 PL, SE 문제에 적용할 계획을 갖고 있다고 했다.

정말 재미있는 발표였다. 다시금 그 사람의 과거 논문을 읽어보게 만드는 힘이 있었다. 나와 비슷한 연구를 하는 사람이라 그렇기도 했지만, 연구자세, 알찬 내용, 자신감있는 표정과 어조도 한 몫 했다.

Toward Gamification of Proofs about Programs

Sorin Lerner

흥미로운 제목에 이끌려서 찾게된 강연이다. 발표자의 과거 논문 목록을 보면 주로 Coq 을 이용하여 엄밀하게 검증을 하던 사람임을 알수 있다. 이런 경험을 바탕으로, 증명을 게임화 하는 몇가지 연구를 소개하였다. 그 중 하나는 이미 인간-컴퓨터간의 소통을 다루는 CHI 학회에 발표되었는데, 프로그램에 관한 증명을 블럭쌓기로 묘사한 시스템이었다.

게임화에는 두 가지 방향이 있다고 생각한다. 하나는 이 연구처럼 어려운 개념을 쉽게 설명하기 위한 교육적 목적이고, 다른 하나는 기계가 풀기 어려운 문제를 인간이 게임을 즐기는 과정에서 자연스럽게 풀리게 하는 것이다¹. 사실 나는 두번째 방향을 기대하고 강연을 들었는데, 발표자는 자신의 목적이 교육이라고 해서 약간 아쉬웠다. 우리 분야의 자동 증명, 자동 검증, 정적 분석, 테스팅 등 기계적으로는 풀기 어려운 학술적인 문제를 재미있게 게임화 할 수만 있다면, 대한민국은 PC 방이라는 상아탑 덕택에 많은 연구 성과를 쏟아낼 수 있을 텐데. 이 강연의 본래 의도는 아니었지만, 기계학습의 열풍 때문에 잠시 잊고 있었던 휴먼컴퓨팅을 다시 떠올리게하고 생각할 거리를 많이 던져준 강연이었다.

PROSE: Programming using Examples

Sumit Gulwani

프로그램 합성(program synthesis) 분야에서 가장 선두를 달리고 있는 Sumit Gulwani 의 강연이다. 프로그램 합성과 Gulwani 의 연구를 자세히 모르는 사람도 MS 엑셀의 자동채우기 기능²을 보여주면 누구나 이해할 정도로, 널리 쓰이는 최첨단 기술을 만들어 가는 사람이다. 즉, 사람이 준 몇 가지 예제를 보고 잘 일반화해서 사용자의 의도에 맞는 프로그램을 만드는 합성 기술을 만드는 것이 그의 목표이다.

대부분 프로그램 합성 연구가 그렇듯이, 굉장히 구체적인 목표를 잡는 것이 핵심인것 같다는 것을 다시 한 번 상기시켜 주는 강연이었다. 일반적으로 프로그램을 자동 생성한다는 것은 쉽지 않은 일이기 때문에 정의된 목표 뿐만 아니라 사용자의 패턴도 잘 알아야 좋은 합성 기술이 탄생한다. 특히, Gulwani 는 주로 엑셀을 기반으로 연구하기 때문에 엑셀 사용자들이 질문 올리는 게시판을 주야장천 살펴본다고 한다. 거기서 사람들이 자주 올리는 질문, 쉽게 틀리는 것들이 곧 새로운 연구주제가 된다고.

요즘 여러 가지 합성 기술이 각기 다른 색깔을 뽐내면서 등장하고 있는데, 이를 한데 아우르는 틀은 가능할까? 이번 PLDI 에도 합성 논문이 여러편이 등장했다. 각자 다른 목표를 가지고 다른 스타일로 접근을 하는데, 이를 한데 아우르는 틀이 있을 것 인가를 가지고 그 날 밤 술잔을 앞에 두고 몇몇이 모여 토론

¹ 대표적인 예로는 단백질의 구조를 분석하기 위한 foldit, 뉴런의 구조를 분석하기 위한 eyewire 등이 있다.

² 가장 단순한 예로, 각 행에 2,4,6을 사람이 입력하고 드래그 하면 8, 10, 12 ... 을 자동으로 채워주는 기능.

을 하였다. 옛날 옛적, 각기 다른 목표를 가진 분석 기술이 많이 등장했지만 나중에 요약 해석 틀 (abstract interpretation framework) 에 비추어보니 다 공통적인 특성이 보였던 것처럼 합성에도 그런 틀이 있을지. 아직은 합성 분야가 생긴지 얼마 되지 않았고 깊이 공부해보지 않아서 잘 모르겠지만, 현재 나는 좀 부정적이다. 프로그램 분석 같은 경우에는, 함수, 고정점, 단조, 연속 등 수학의 오랜 역사동안 차곡차곡 쌓여온 좋은 개념을 잘 이용할수 있었다. 반면, 사용자의 의도를 해석해야하는 합성 분야에서 그런 좋은 성질과 개념이 존재할까? 관심을 갖고 지켜볼 일이다.

인상깊은 논문

Data-Driven Precondition Inference with Learned Features

Saswat Padhi, Rahul Sharma, Todd Millstein

학회에 참가 하기 직전, 요즘 하고 있는 연구 내용과 비슷한 제목이 PLDI 에 발표된 것을 보고 궁금해하며 발표를 기다렸다. 기계학습을 이용하여 정적분석의 성능을 높이는 연구를 하면서 늘 문제가 되는 것이 특징 (feature) 를 만드는 것이다. 기계가 잘 학습하도록 학습 데이터를 잘 설명하는 특징을 제공해야하는데 이것을 늘 사람이 만들어야 하는 것이 문제였다. 그러다보니 골치도 아팠고 남들에게 기술의 일반성을 이야기하기도 좀 난감했다. 그래서 특징을 자동으로 생성하는 연구를 하고 있는 찰나에, 이 논문에서도 특징을 자동으로 배운다니 아주 관심이 가는 논문이었다.

다행인지 불행인지 발표를 듣고나니 우리가 하고 있는 일과 큰 관계는 없어보였다. 우선 이 사람들의 목표는 함수의 전제조건 (pre-condition) 을 자동으로 찾아내는 것이었다. 그 전제 조건은 함수 입력이 만족해야할 조건을 부울식으로 표현한 것이다. 함수의 올바른 입력과 틀린 입력 데이터가 많이 제공되면 이를 바탕으로 전제조건을 추론하는 것에서 출발한다. 보통은 시작하기 전에 미리 정해둔 재료를 이리저리 조합해보면서 최대한 일반적인 전제 조건을 만들려고 노력한다. 그 재료는 간단한 부울 식이다. 예를 들면, $f(x,y)$ 라는 함수가 있을때 $(x < 0)$, $(x > 0)$, $(y < 0)$, $(y > 0)$ 같은 기본 재료를 이리 저리 조합해가면서 전제 조건을 만드는 것이다. 그런데 문제가 복잡한 경우에 미리 정해둔 재료로는 도저히 현재 가지고 있는 예제를 다 만족시키지 못하는 경우가 있다. 이 때 새로운 재료가 필요한데 이를 프로그램 합성 기술로 만들어 내는 것이 핵심이다. 이 논문에서는 숫자와 문자열을 다루는 프로그램을 대상으로 하기 때문에 $+$, $-$, $length$ 같은 기본적인 숫자, 문자열 연산자와 입력 데이터를 바탕으로 필요한 조건을 합성해 낸다.

“합성”이라는 단어가 주는 마법같은 (무에서 유를 창조할 것 같은) 느낌 때문에 엄청난 것이라 생각할 때도 있었는데, 사실 따져 보면 공간 탐색 문제를 푸는 것이라 할 수 있겠다. 다른 분야도 마찬가지로일테지만, 합성 분야 논문을 보면 유독 실용적인 문제를 잘 찾는 것이 핵심인 것 같다. 제각기 실용적인 문제를 귀신같이 풀어내는 것을 보면 신기하기는 하지만, 그 해결 방식에 공통점이 별로 없어 보여서 서로 기술을 쌓아간다는 느낌이 들지가 않아 조금 아쉽다.

Cartesian Hoare Logic for Verifying k-Safety Properties

Marcelo Sousa, Isil Dillig

프로그램의 올바름을 따질 때에는, 프로그램이 정의되지 않은 행동을 하지 않는지 안전성 (safety) 을 따지는 것을 넘어서, 프로그램이 원하는 요구사항을 잘 만족하는지 기능성 (functionality) 을 검사하는 것도 필요하다. 일반적으로 기능성을 따지는 것은 안전성을 따지는 것 보다 더 어려운 문제이다. 보통 안전성을 따질 때는 프로그램의 어떤 한 실행이 이상한 행동을 하는 지를 검사하지만, 기능성을 따질 땐 프로그램의 여러 실행이 어떤 상관 관계를 갖는지를 검사해야한다. 예를 들어, 어떤 함수 f 의 실행이 늘 결정적 (deterministic) 인지를 검사하려면, $x_1 = x_2$ 를 만족하는 두 입력 x_1, x_2 를 각각 실행한 $f(x_1)$ 와 $f(x_2)$ 가 서로 다른 값을 낼 수 있는지 검사해야한다. 이렇게 기능성 오류를 검사할 때 함께 따지는 k 개 실행의 성질을 일반적으로 k -safety 라고 한다. 이 논문에서는 이러한 k -safety 를 검증할수 있는 논리 시스템을 고안하였다. 호어 논리 (Hoare logic) 의 확장판인데 의외로 아이디어는 간단하다. 기본적인 호어 논리에서는 전제 조건 (pre-condition) 과 사후 조건 (post-condition) 에 프로그램 실행 하나에 대한 이야기가 담기지만 이 논문에서 제시하는 Cartesian Hoare Logic 에서는 k 개 실행의 관계에 관한 이야기를 담는 것이다. 예를 들면, f 가 결정적인지를 검사하려면

$$\{x_1 = x_2\} f(x) \{ret_1 = ret_2\}$$

이처럼 일종의 관계분석을 하는 것이다.

기능성 오류를 표현하는 k -safety 라는 개념을 이 논문을 통해 처음 알았는데, 흥미로운 아이디어이다. 이 문제를 풀기 위한 Cartesian Hoare Logic 은 목적은 다르지만 개념적으로 관계분석과 매우 비슷하는데, 더 실용적인 k -safety 분석기를 팔각관계분석 (octagon relational analysis) 등으로 만들 수 있지 않을까 하는 생각이 얼핏 들었다.

Statistical Similarity of Binaries

Yaniv David, Nimrod Partush, Eran Yahav

이 연구는 굉장히 실용적인 동기와 해결책이 돋보인다. 보통 널리 쓰이는 중요한 소프트웨어에서 취약점이 발견되면, 보안회사들은 그 해당 취약점이 담긴 소프트웨어의 바이너리 샘플을 가지고 각 가정이나 회사에 그 문제가 있는 소프트웨어가 있는지를 검사한다. 하지만 이 때 각 현장마다 쓰는 OS 가 다르고, 컴파일러의 종류, 버전이 다르며 해당 소프트웨어의 버전도 약간씩 다른 것이 문제이다. 이런 경우에 같은 소스 코드에서 컴파일된 바이너리지만 생김새가 매우 달라서 찾아내기가 힘들다. 이 연구에서는 이렇게 복잡한 상황에서도 바이너리로 작성된 프로그램의 유사도를 잘 측정하는 방법을 제시하였다.

이 논문의 전반적인 문제의식에 크게 공감한다. 우리는 소스 코드를 분석하여 개발자에게 도움을 주는 방법을 계속 연구하고 있지만, 세상엔 소스 코드를 얻을 수 없는 상황이 많다. 날마다 보고되는 보안 취약점을 살펴보려 했을 때도 대부분 소스코드가 공개되어 있지 않은 소프트웨어였다. 이러한 상황에서 우리가 할 수 있는 것은 무엇인지 심각하게 고민해보는 계기가 되었다.

From Datalog to Flix: A Declarative Language for Fixed Points on Lattices

Magnus Madsen, Ming-Ho Yee, Ondřej Lhoták

이 연구에서는 고차원 언어로 간편하게 기술한 프로그램의 분석 명세에서 분석기를 자동으로 만들어주는 시스템을 제안하였다. 제목에서도 알 수 있듯이, 기존에 Datalog 라는 언어의 표현력을 확장한 것이 Flix 이다. Datalog 프로그램은 여러 규칙 (rule) 으로 구성되고, 목표는 그 규칙을 만족하면서 가장 작은 모델을 찾는 것이다. 규칙을 기술하기가 쉽고 모델을 효율적으로 찾을 수 있기 때문에, 프로그램에서 뽑은 방정식의 최소 고정점을 찾을 때 (즉, 정적 분석을 할 때) 만들때 많이 사용한다. 하지만 Datalog 는 제약이 있어서 비교적 간단한 분석을 기술하는데 밖에 사용하지 못하는데, 이 연구는 그 제약을 어느정도 극복하여 좀 더 일반적인 분석을 할 수 있게 만든 것이 핵심이다. Datalog 는 모든 규칙을 명시적으로 하나씩 나열해야 했기 때문에 무한한 규칙을 기술할 수 없다. 따라서 pointer 분석 등 유한한 대상을 다루는 분석에 주로 적용을 했다. 반면 Flix 는 무한한 래티스 (lattice) 와 함수를 기술할 수 있기 때문에 상수 분석, 정수 구간 분석도 기술할 수 있다.

사람이 고차원 언어로 기술한 정적 분석을 자동으로 생성하는 이야기는 그다지 새로운 것이 아니지만 실제로 돌아가는 시스템을 만들고 공개한 것이 큰 기여가 아닌가 생각한다. 간단한 포인터 분석 뿐만 아니라 정수 값 분석도 가능하다고 하니, Sparrow 로 연구하고 있는 나로서는 Flix 로 만든 분석기가 사람이 손수 다 만든 것에 비해 얼마나 성능이 뛰어날지 궁금하다. 논문에 있는 실험 결과는 비교적 간단한 분석기에 관한 이야기여서 조금 아쉽다. 또한 이 연구를 보면 우리 연구실에서 개발하고 있는 분석기+검산기 자동 생성 시스템도 완성되면 좋은 평가를 받을 수 있을 것 같다.

학회 즐기기

이번 학회가 열린 산타바바라는 미국 캘리포니아에 있는 작은 휴양도시이다. 찾아보니 스페인 이민자들이 건설한 도시라고 하는데 수많은 작은 집이 온 산을 뒤덮고 있는 달동네가 인상적이었다. 동네가 작다 보니 시내까지 걸어다닐 수 있었다. 날씨는 역시 캘리포니아다 시원 정도로 맑고 화창했다. 점심은 놀이 테이블에 준비해주었는데, 날씨 덕분에 그 맛이 배가 되었다.



사람들을 만나서 교류하는 것은 즐거운 일이다. 이 분야에서 학생으로 공부한지가 그래도 제법 되다 보니 여기저기서 안면이 있는 사람들이 보였다. 다른 학회에서 만났던 사람들도 있었고, 서울대를 방문했던 사람들도 있었다. 부끄러운 일이지만 예전에는 늘, 더 연구 경험이 쌓이면, 더 영어가 유창해지면 활발히 교류할 수 있을 것이라고 생각했다. 허나 늘 그렇듯 모든 채비를 다 갖춘채로 시작할 수 있는 행운은 우리에게 없다. 또한 수 년전보다 영어실력과 연구 경험이 조금 더 늘었다고 해서 딱히 달라진것도 없었다. 결국 학회에서 사람들과 교류하는 것은 습관이다. 그리고 습관은 대부분 한방에 고치지 않으면 안된다. 그래서 이번에는 기회가 되는 대로 눈 마주치는 사람에게 다가가서 자기 소개를 하고 연구 소개를 한 후 포스터 발표를 하니 오라고 이야기해 보았다. 시큰둥한 사람도 있었지만 대부분은 반갑게 자기 이야기도 해주었다. 간혹 말이 너무 많아서 지루한 사람도 있었지만 영어 공부 하는 셈치고 유심히 들어보았다. 포스터에 찾아온 이도 있었다. 습관을 쉽게 고치게 해준 고마운 사람들이다. 또한 소프트웨어무결점 연구센터에서 인연을 맺은 류석영 교수님과 카이스트 학생들, 같이 연구하는 고려대 오학주 교수님과 고려대 학생들, 우리 연구실에서 공부하다가 유학을 떠난 원태형과도 오랜만에 만나서 알찬 이야기를 나누었다.



마무리

학회를 참가하면 늘 여러가지 생각이 든다. 어떤 사람들과 어떤 이야기를 나눌지 설레는 마음으로 가서, 대가들의 연구를 보면서 입이 벌어지기도 하다가, 살펴보면 또 의외로 다들 별거 아니다 싶은 마음도 들다가, 결국 나는 그래서 무얼 해야하나로 귀결된다. 다행히도 좌절보다는 활기와 오기를 가득 안고 돌아

온 학회였다. 앞으로 하고 싶은 일도 더 명확해지고 몰랐던 것도 많이 알게 되었다. 이번 PLDI 에 논문도 없었는데 경진대회에 참가하라고 권유해주시고 지원해주신 이광근 교수님께 감사드린다.