실용적인 파이썬 분석기 만들기



이석현, 오학주
소프트웨어 분석 연구실, 고려대학교

2026.8.26.

2026 SIGPL 여름학교 @한국문화연수원, 공주

정말 중요한 파이썬

- PYPL(2026.08.)기준 전세계 50%가 관심을 가지는 언어

Worldwide, Aug 2026 :

Rank	Change	Language	Share	1-year trend
1		Python	50.44 %	+20.2 %
2		Java	12.8 %	-2.0 %
3		C/C++	8.57 %	-1.2 %
4	↑	R	4.18 %	-1.3 %
5	↓	JavaScript	3.52 %	-3.4 %
6	↑↑↑↑↑↑	Objective-C	3.01 %	+0.9 %
7		PHP	2.06 %	-1.4 %
8	↑	Rust	1.86 %	-1.0 %
9	↓↓↓	C#	1.74 %	-3.0 %
10	↓↓	Swift	1.68 %	-1.2 %

정말 중요한 파이썬

- PYPL(2026.08.)기준 전세계 50%가 관심을 가지는 언어
- 중요한 분야에서도 충분히 활용되고 있음





보안 및 인증





인프라 관리



금융



정밀 기계 제어

정말 중요한 파이썬

- PYPL(2026.08.)기준 전세계 50%가 관심을 가지는 언어
- 중요한 분야에서도 충분히 활용되고 있음
- 배포하기 전에 **오류를 미리 발견하는 것**이 매우 중요
- 오류를 미리 발견하는 방법
  - ***실행하기 전에 검사(정적 분석)***
  - 마구잡이로 실행해서 검사(테스팅)

파이썬에서 정적 분석, 뭐가 어려울까?

1. 문법에 숨겨진 동작이 너무 많음(SIGPL'25 여름학교)
   - "+"가 얼마나 어려운지
   - Mir(미르) 파이썬을 분석하기 위한 중간 표현



Table 2. Fidelity evaluation results of Mir.

Benchmark	Total	Applicable	Compiled	Executed	Correct
Mir benchmark	266	266	266	266	266
MBPP	974	619	619	619	619
HumanEval	163	116	116	116	116
Total	1,403	1,001	1,001	1,001	1,001

Table 3. Soundness and completeness of MirCG and PyCG on the benchmarks with ground-truth call graphs.

		MirCG			PyCG [45]		
Benchmark	Total	Sound	Complete	Time (s)	Sound	Complete	Time (s)
MirCG benchmark	100	**87**	77	5.2	15	77	4.9
Micro-benchmark	112	**107**	106	6.7	46	106	5.9
Total	212	**194 (91.5%)**	183 (86.3%)	11.9	61 (28.8%)	183 (86.3%)	10.8

Table 5. Soundness and completeness of MirCheck and Nagini. **Red** marks unsound verdicts (accepting an incorrect specification) and **orange** marks incomplete verdicts (rejecting a correct specification).

		MirCheck				Nagini [7]			
Spec. Type	Total	Accept	Reject	Inconc.	Time (s)	Accept	Reject	Inconc.	Time (s)
Correct	25	25	0	0	4.4	16	5	4	4.5
Incorrect	25	0	25	0	4.4	4	17	4	4.5

파이썬에서 정적 분석, 뭐가 어려울까?

2. 복잡한 동작(확장성)
   - 아주 작은 조각을 컴파일 해도 그 안에는 십수개의 분기가 있음
   - 큰 실제 프로젝트를 분석하기에는 아직 너무 느림
   - 500KB 파이썬 코드를 MIR로 컴파일 하면 40MB가 됨

3. 실행하기 전에는 절대 모르는 부분(정확성)
   - 가능한(의도와 다른) 실행이 너무 많음
   - 모든 실행을 고려하면 알람이 지나치게 많아짐
   - 평균 알람의 수: 1000줄 당 500개

실용적인 파이썬 분석기

- 목표: **실용적인** 파이썬 분석기
  - 적절한 시간과 자원을 쓰고,
  - 오류는 많이 찾으면서,
  - 알람의 수는 적은

