JavaScript 웹 앱 정적 분석을 위해 다양한 모델링을 조화롭게 합치기

배성경 imai0917@kaist.ac.kr

박지혁 jhpark0223@kaist.ac.kr

류석영 sryu@cs.kaist.ac.kr

연구 동기

JavaScript 웹 앱들은 다양한 API를 사용하는 경우가 많다. 정적 분석을 제대로 하기 위해서는 이런 API들을 모델링해야 한다.

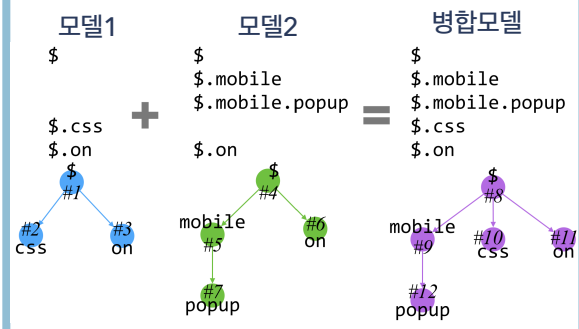
장단점을 가진 다양한 모델링 기법이 존재한다.

수동 모델링	자동 모델링
– 노동 집약적	– IDL로 쓰인 스펙 필요
– 분석기에 붙박이	– 자세한 모델링 불가능
+ 자세한 모델링 가능	+ 자동으로 모델링

각각의 장점을 살리면서 조화롭게 병합하고 싶다.

수동: 자세한 의미 + 자동: 함수 인자의 타입 / 개수 검사, 콜백 함수 호출

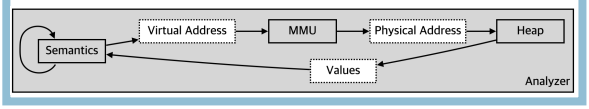
힙 병합



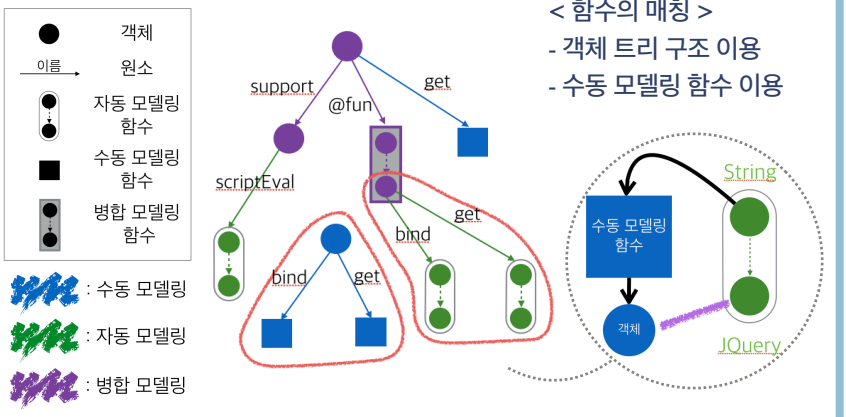
M M U 가상 ↕ 물리

#1 ↦ #8, #2 ↦ #10, #3 ↦ #11,
#4 ↦ #8, #5 ↦ #9, #6 ↦ #11, #7 ↦ #12

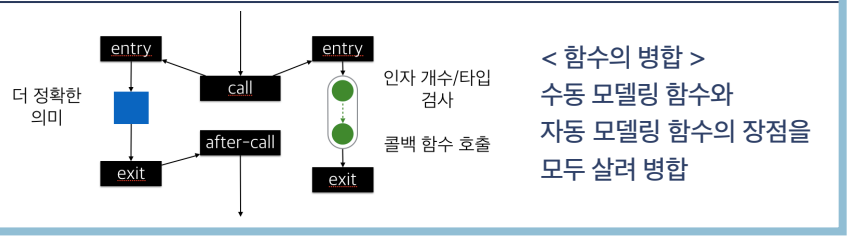
- 각각의 모델은 각자 힙을 모델링 (가상 주소)
- 병합된 힙에서는 새로운 주소 할당 (물리 주소)
- 가상주소로 접근하면 물리주소로 맵핑 (MMU)



함수 병합



< 함수의 매칭 >
- 객체 트리 구조 이용
- 수동 모델링 함수 이용

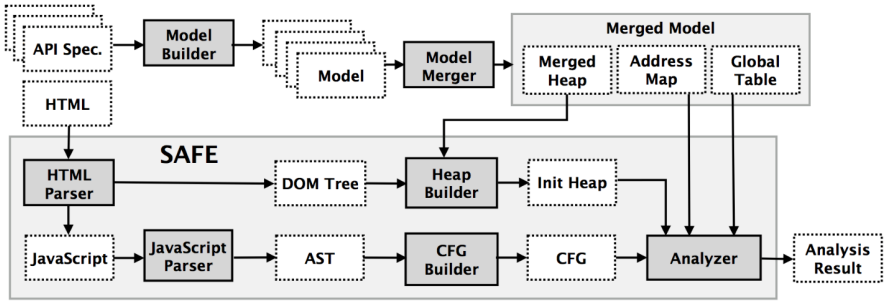


< 함수의 병합 >
수동 모델링 함수와 자동 모델링 함수의 장점을 모두 살려 병합

병합의 특성

- 힙 병합은 올바르다
병합 후 객체가 가지는 원소의 집합은 기존 객체들이 가지는 원소들의 합집합과 동일

- 합수 병합은 효과적이다
동일한 의미를 가지는 함수를 매칭하고 각 모델링 함수의 장점을 살려 병합



결과

App.	LOC JS	LOC HTML	User function coverage Man.	User function coverage Auto.	User function coverage All	API function coverage Man.	API function coverage Auto.	API function coverage All	Time (sec.) Man.	Time (sec.) Auto.	Time (sec.) All	Issues Man.	Issues Auto.	Issues All
App1	52	19	3/5	3/5	3/5	2/7	2/7	2/7	4	5	6	0	1	1
App2	972	69	47/126	8/126	47/126	17/30	2/31	17/31	27	8	57	3	10	48
App3	1339	211	118/160	9/160	122/160	42/46	2/51	42/51	1027	11	2240	14	7	84
App4	206	76	25/27	27/27	27/27	24/26	24/28	24/28	122	122	334	4	31	35
App5	105	29	12/20	15/20	15/20	8/8	8/8	8/8	7	6	13	1	8	10
App6	608	30	31/96	22/96	34/96	17/27	16/27	20/27	18	9	30	0	22	26
App7	122	41	20/23	22/23	22/23	10/12	9/13	11/13	19	14	48	1	24	25
App8	145	46	7/13	8/13	8/13	9/13	9/13	9/13	11	16	24	2	9	11
App9	7883	96	6/10	8/10	8/10	8/11	10/11	10/11	8	8	14	1	22	22

< 병합 모델을 통한 분석 결과 >
Tizen 샘플 웹 어플리케이션 23가지 중 9가지를 선별하여 분석 (선별 기준 - API의 사용 유무)

- 기존 모델들이 분석하는 함수 포함 / 더 많은 함수 분석
- 기존 모델들이 검출한 이슈 포함 / 이슈 검출 능력 향상