Capgen vs SimFix

1. 각각의 기술이 커버하는 pattern 인지 여부
2. Capgen의 주요 AST 매칭

* Ancestor 차수가 높을수록 (기준점은 두 번 이상의 degree상승시)
* 의심 node와 같은 차수의 sibling node들이 다양하게 존재 하는지

1. SimFix

* 버그 stmt가 충분한 snippet을 가지는지
* Snippet 내의 variable 들이 충분히 다른 타입들의 expression으로 사용되는지 (2개 이상)

1. 구체적이고 수학적인 수식을 제안하는 것은 전체 AST를 활용하는 기술에 대한 포괄적인 접근과 거리가 있고 나의 관심사가 아님.

두 기술이 옳은 패치를 생성하는 모든 케이스를 위의 기준으로 예측하여 두 기술중 해당 bug AST구조를 보고 고칠 수 있는 기술을 예측 or 둘다 고친다면 더 효율적으로(시간적인) 고칠 수 있는지 예측

chart1 : Capgen (o)

Chart7 : simfix 충분히 snippet존재하며 variable 타입들이 존재(o)

chart8 : Capgen : 충분한 snippet 수가 없다. (o)

chart 11 : capgen variable 타입들 x 충분히 sibling들 존재 (o)

chart 20 : SimFix 여러 타입의 variable들이 사용되고 있음 (o)

chart 24 : Capgen : 충분한 snippet 수가 없다 (o)

lang 6 : Capgen : 충분한 snippet이 없으며 계층적인 구조가 잘 이루어져 있음. (o)

lang 26 : Capgen : 충분한 snippet 담지 못하며 계층적인 구조와 sibling은 충분함. (o)

lang 27 : SimFix : 충분히 snippet이 존재하며 같은 타입의 sibling 반복 (o)

lang 33 : Capgen : snippet이 부족하며 계층적인 구조 존재 (x)

lang 39 : SimFix : snippet충분하며 2개 이상의 변수들이 반복되고 있다. (variable에 대한 사용 추정 용이) (o)

lang 43 : SimFix : snippet 충분하고 2개 이상의 변수들이 반복되고 있다. (o)

lang 57 : Capgen 충분한 snippet도 없으며 2개이상의 변수들이 타입 파악을 위해 사용되지 않음.(o)

lang 58 : SimFix snippet이 충분하고 변수들이 여러 타입에 사용되고 있다 (o)

lang 59 : Capgen 계층적으로 구조가 뚜렷하며 (ancestor degree 2이상) (o)

lang 60 : SimFix snippet 충분하고 변수가 다양한 타입으로 사용됨. (o)

math 5 : capgen snippet이 충분한 변수 변화를 가지지 못함. (X)

math 30 : capgen 계층적인 구조가 뚜렷히 존재하며 sibiling들이 충분히 많음. (o)

math 33 : capgen 계층적인 구조 뚜렷하고 sibling 등도 충분히 존재 (x)

math 41 : capgen 위와 같음 (x)

math 50 : simfix variable이 다양한 타입으로 사용됨 (o)

math 53 : SimFIx snippet이 충분하고 변수가 다양한 type으로 사용됨 (o)

math 57 : SimFix snippet 내의 variable등이 다양한 타입별로 사용되고 snippet의 크기 충분 sibling의 정보가 약함. (o)

math 58 : SimFix(x).

math 59 : simfix snippet 내의 variable이 다양한 타입으로 제공 ???, capgen도 충분히고려(X, 근소한 차이)

math 63 : simfix snippet 내의 variable에 다양한 타입으로 제공 (o,근소한 차이)

math 70 : simFIx variable 타입들 산출 가능(o)

math 71 : capgen 계층적 구조가 깊고 snippet이 부족하며 variable 타입들이 충분히 제공되지 않음 (x)

math 75 : SimFix Variable 타입에 대한 정보 제공(o)

math 79 : simfix snippet 충분하고 variable 타입들이 충분히 제공됨(o)

math 80 : SimFix 계층적 구조가 깊으며 snippet 내의 variable 타입들이 충분히 제공(x)

math 85 : capgen 계층적 구조 깊고 snippet 내 variable타입들이 충분히 제공 x

math 90 : simfix snippet 충분하며 variable 타입들이 여럿 사용됨.(o)

prediction rate 24/33

capgen : 9/12

simfix : 15/21