Homework 3
SNU 4541.664A

Kwangkeun Yi

- 이번 숙제의 목적은 다양한 문제에서 정적분석이 유용한 경우에 맞추어 정적분석기를 디자인하는 것이다.
- 점프명령어 때문에 프로그램 실행을 드러내는 전이과정(transitional) 스타일로 분석기를 디자인하게 될 것이다.
- 디자인 기한: 11/15(화) 15:30 프린트제출.

아래의 언어를 생각하자. 지난 숙제에서 대상으로 한 언어를 살짝 확장한 언어이다. 정수입력을 받고, 변수의 주소도 값으로 다룰 수 있도록 했고, 타겟이 실행중에 계산되는 점프명령이 있다.

$$
\begin{array}{rcl}
C & \to & x \texttt{ := } E \mid \texttt{*}x \texttt{ := } E \\
& \mid & C \texttt{ ; } C \\
& \mid & \texttt{if } E\ C\ C \\
& \mid & \texttt{repeat } C\ E \\
& \mid & \texttt{goto } E \\
E & \to & n \quad (n \in \mathbb{Z}) \\
& \mid & E \texttt{ + } E \\
& \mid & E \texttt{ * } E \\
& \mid & \texttt{-} E \\
& \mid & E \texttt{ < } E \\
& \mid & x \mid \texttt{*}x \mid \texttt{\&}x \\
& \mid & \texttt{readInt} \quad \text{(정수입력)}
\end{array}
$$

점프명령어 "goto E"에서 식 E를 계산하면 점프할 명령어(C)의 번호가 된다. 그 명령어로 점프한다. 프로그램의 모든 명령어마다 고유의 자연수가 명령어 번호로

붙는다고 가정한다. 고유 번호 ℓ이 붙은 명령어 $\ell : C$는 다음과 같이 만들어진다:

$$\begin{array}{rcl} C & \rightarrow & \ell : C' \\ C' & \rightarrow & x \texttt{ := } E \mid \texttt{*}x \texttt{ := } E \\ & \mid & C \texttt{ ; } C \\ & \mid & \texttt{if } E\ C\ C \\ & \mid & \texttt{repeat } C\ E \\ & \mid & \texttt{goto } E \end{array}$$

위의 언어로 짜여진 프로그램을 정적분석하는 다음의 두 가지 분석기를 디자인한 결과를 제출하라.

Exercise 1 (100pts) "daVinci 코드 검증기"

다빈치가 심혈을 기울여 위의 언어로 개발한 소프트웨어. 그 소스에 다빈치가 원저자임을 식별할 수 있는 고유의 식별 코드 조각이 뿌려져있다고 알려져 있다. 최근에 다빈치 코드의 성질은 다음과 같은 것임이 밝혀졌다:

> 어떤 변수가 있어서, 소프트웨어가 실행중에 그 변수가 가지는 정수값은 항상 1867(다빈치의 탄생년도 1452 와 탄생월일 0415를 더한 값)으로 나누어 나머지가 415(다빈치의 탄생월일)이다.

그러나 주어진 소프트웨어가 다빈치의 진품인지 아닌지를 확인하려면 간단치가 않다. 소프트웨어들을 모든 입력의 경우마다 일일이 돌려보고 그런 성질의 변수가 있는지를 찾아내야 하는데, 이 과정이 비용이 많이 들거나 심지어는 아예 불가능하게 된다. 입력의 경우가 무한하거나 너무 많은 경우도 그렇고, 소프트웨어가 사용하는 많은 변수들의 값들을 실행을 통해서 일일이 추적해야 하는 어려움도 그렇고.

좋은 방법은 정적분석 기술을 이용하는 것이다. 정적분석기술을 이용하면 주어진 소프트웨어가 위와 같은 변수를 가지고 있는지를 자동으로 안전하게 판단할 수 있게 된다. 정수 변수들이 실행중에 가지는 값들이 1867로 나누어 어떤 나머지 값들을 가지는지를 분석하면 된다.

□

예를들어, 여러분의 분석기는 다음의 소스를 분석해서 다빈치의 진품임을 확인해 줄 수 있어야 한다. 변수 xp가 실행중에 가지는 값은 항상 1867로 나누어 415이기 때문이다.

```
port = 196; len = 43; pos = 139;
index = 101; fd = 15; xp := 2282; count = 0;
```

```
i = xp;
repeat
     len = len + port;
     pos = pos * i;
     index = index * port;
     xp = i + 3734;
     fd = fd * port;
     if (sock + 1 < pos) index = index +1
        index = index + (-1)
(pos*10 < port)
```

□

Exercise 2 (100pts) "SW매연 저감장치 검증기"

4차산업혁명도 심각한 매연을 뿜을 것이다. 앞으로는 우리들의 모든 면(믿을만함, 직업적합성, 신용도, 성실성, 창의성, 긍정성, 등등)의 평판(score)이 컴퓨터로 결정되고 그 평판에 기반해서 중요한 결정이 내려지는 사회가 펼쳐질 수 있다. 컴퓨터가 계산하는 평판에 기대어 각자의 미래가 한정지어지는 숨막히는 지옥으로 떨어질. 내가 무엇을 하려고 하던지간에 나는 컴퓨터가 점수매기는 나의 평판에 얽매여진다. 돈을 빌리고 싶어도, 취직을 하고 싶어도, 사업을 하고 싶어도, 동업자를 소개받고 싶어도 내 과거의 모든 일거수일투족이 내 미래의 발전가능성을 결정하는 것이다.

이러면 평판의 양극화가 심각해진다. 좋은 평판을 가진 사람은 계속해서 잘 나갈 수 있지만, 한 번 실수로 평판이 낮아진 전과를 가진 사람은 재기할 수 있는 여지가 없어진다.

이에 2025년 국회는 법률을 제정한다. 평판의 양극화를 막는 법률이다. 공공영역에서 사용되는 모든 평판계산 소프트웨어는 안전장치를 갖춘 것만 허용하는 법률을 제정한다.

안전장치란, 입력값을 통해서 평판계산 알고리즘(소프트웨어)의 맹목적인 질주를 억제할 수 있도록 한 것이다. 입력값이 99를 넘을때 소프트웨어 내의 변수 `liberation`이 1을 가져야하고, `liberation`이 1이 되면 늘하던 평판계산 대신에 80에서 100사이의 값을 평판(결과값)으로 내놓아야 한다. 이런 장치를 통해서 공공기관에서는 법률이 정한 빈도에 맞추어 해당하는 입력값을 넣어줌으로서 평판계산 알고리즘이 종종 무작위로 작동하도록 하는 것이다. 그래서 그동안 평판이 나빴던 사람들에게도 회생의 기회를 운좋게 주도록.

이 법률의 집행은 [SW매연 저감위원회]가 인증을 맡는다. 그 위원회에서는 제출된 소프트웨어의 인증여부를 결정한다. 그런 장치가 제대로 갖춰졌는지를 확인하고.

우리의 목표는 그 위원회가 사용할 정적분석도구를 제작하는 것이다. 입력값이 99을 넘을 때 변수 `liberation`이 1 되고 평판값 변수 `score`가 80에서 100사이의 값이되는 지를 확인하는 분석기이다. □