

SNU 4541.664A Program Analysis

Spring 2005

Note 10

Prof. Kwangkeun Yi

요약해석 디자인과 구현의 예

변수가 있는 정수식 프로그램의 요약해석
명령형 언어 프로그램의 요약해석

변수가 있는 정수식 프로그램의 요약해석

$$\begin{array}{l}
 E \rightarrow n \quad (n \in \mathbb{Z}) \\
 | \\
 | \quad E + E \\
 | \\
 | \quad - E \\
 | \\
 | \quad x \quad \text{변수} \\
 | \\
 | \quad \text{let } x = E_1 \text{ in } E_2 \quad \text{지역 변수}
 \end{array}$$

의미공간

$$\sigma \in Env = Var \xrightarrow{\text{fin}} 2^{\mathbb{Z}}$$

의미함수

$$\mathcal{V} : Exp \rightarrow Env \rightarrow 2^{\mathbb{Z}}$$

$$\mathcal{V} n \sigma = \{n\}$$

$$\mathcal{V} x \sigma = \sigma x$$

$$\mathcal{V} E_1 + E_2 \sigma = \{z_1 + z_2 \mid z_i \in \mathcal{V} E_i \sigma\}$$

$$\mathcal{V} - E \sigma = \{-z \mid z \in \mathcal{V} E \sigma\}$$

$$\mathcal{V} \text{let } x = E_1 \text{ in } E_2 \sigma = \mathcal{V} E_2 \sigma \{x \mapsto \mathcal{V} E_1 \sigma\}$$

모든 식 E 의 의미는 조립식으로 $\mathcal{V} E$, 즉 상수함수 $\lambda x. \mathcal{V} E$ 의 최소고정점.

요약된 의미함수

$$\hat{V} : Exp \rightarrow Env \rightarrow \hat{\mathbb{Z}}$$

의 의미공간 사이의 갈로아 연결

$$Env \rightarrow 2^{\mathbb{Z}} \xrightleftharpoons[\alpha]{\gamma} \hat{Env} \rightarrow \hat{\mathbb{Z}}$$

은 조립식으로, $Env \xrightleftharpoons[\alpha_1]{\gamma_1} \hat{Env}$ 와 $2^{\mathbb{Z}} \xrightleftharpoons[\alpha_2]{\gamma_2} \hat{\mathbb{Z}}$ 을 가지고:

$$\alpha f = \alpha_2 \circ f \circ \gamma_1$$

$$\gamma \hat{f} = \gamma_2 \circ \hat{f} \circ \alpha_1$$

$$\hat{\mathcal{V}} n \hat{\sigma} = \alpha_2 \{n\}$$

$$\hat{\mathcal{V}} E_1 + E_2 \hat{\sigma} = (\hat{\mathcal{V}} E_1 \hat{\sigma}) \hat{+} (\hat{\mathcal{V}} E_2 \hat{\sigma})$$

$$\hat{\mathcal{V}} - E \hat{\sigma} = \hat{-} (\hat{\mathcal{V}} E \hat{\sigma})$$

$$\hat{\mathcal{V}} \text{let } x = E_1 \text{ in } E_2 \hat{\sigma} = \hat{\mathcal{V}} E_2 \hat{\sigma} \{x \mapsto \hat{\mathcal{V}} E_1 \hat{\sigma}\}$$

정수식 E 의 요약 의미는 조립식: $\hat{\mathcal{V}} E$. 확인할 것: 모든 정수식 E 에 대해서

$$\alpha(\mathcal{V} E) \sqsubseteq \hat{\mathcal{V}} E$$

구현

프로그램 E 에 대해서, $\hat{V} E \in \hat{Env} \rightarrow \hat{Z}$ 의 계산:

- \hat{Env} 의 원소가 유한하다면?
모든 $\hat{\sigma} \in \hat{Env}$ 에 대해서 $\hat{V} E \hat{\sigma}$ 를 각각 계산.
 $\hat{V} E \hat{\sigma}$ 의 계산은 항상 끝남 (E 에 대한 조립).
- \hat{Env} 의 원소가 무한히 많다면? 혹은, 유한하지만 모든 경우가 필요 없다면?
 - 하나의 요약 환경 $\hat{\sigma}_0$ 으로 초기 환경을 모두 포섭시키고, $\hat{V} E \hat{\sigma}_0$ 을 계산.
 - 반드시 하나의 요약환경일 필요 없음.

구현 예1

자유변수가 없는 프로그램

$$E = (\text{let } x = 1 \text{ in } (\text{let } y = 2 \text{ in } (\text{let } x = 3 \text{ in } x + y))) + x$$

$$\begin{aligned} \hat{V} E \hat{\sigma} &= \hat{V} E_2 \hat{\sigma} \{x \mapsto \hat{V} 1 \hat{\sigma}\} \\ &= \hat{V} E_2 \hat{\sigma} \{x \mapsto 1\} \\ &= \dots \end{aligned}$$

- $\hat{V} E$ 는 상수함수, $\hat{\sigma}$ 와 상관없는.
- $\hat{V} E$ 는 E 의 구조를 타고 조립. 유한시간에 계산가능.

구현 예2

자유변수(입력)가 있는 프로그램

$$E = (\text{let } y = 2 \text{ in } (\text{let } x = 3 \text{ in } x + y)) + x$$

$$\begin{aligned} \hat{\mathcal{V}} E \hat{\sigma} &= (\hat{\mathcal{V}} E_1 \hat{\sigma}) + (\hat{\mathcal{V}} x \hat{\sigma}) \\ &= \dots \end{aligned}$$

함수 $\hat{\mathcal{V}} E$ 는 E 의 구조를 타고 조립식으로 정의되지만, 어떤 함수일까?

- $\hat{\mathcal{V}} E \{x \mapsto \top\}$ 만 계산
- 혹은 각 경우를 모두 계산:

$$\{\hat{\mathcal{V}} E \{x \mapsto \perp\}, \hat{\mathcal{V}} E \{x \mapsto 0\}, \hat{\mathcal{V}} E \{x \mapsto +\}, \dots\}$$

명령형 언어 프로그램의 요약해석

$$\begin{aligned}
 C &\rightarrow \text{skip} \\
 &| x := E \\
 &| \text{if } E \ C \ C \\
 &| C ; C \\
 &| \text{while } E \ \text{do } C \\
 E &\rightarrow n \quad (n \in \mathbb{Z}) \\
 &| x \\
 &| E + E \ | \ E < E
 \end{aligned}$$

의미공간은

$$\text{Memory} = \text{Loc} \xrightarrow{\text{fin}} \text{Value}$$

$$\text{Value} = 2^{\mathbb{Z}} + 2^{\mathbb{B}}$$

세 갈래 길

- 프로그램 C 의 의미를 조립식으로 (최소고정점)
- 프로그램 C 의 의미를 방정식의 해로 (최소고정점)
- 프로그램 C 의 의미를 기계상태의 전이과정으로 (최소고정점)

C 의 의미를 방정식의 해로

일단, 프로그램 C 의 의미를 다음의 함수로 정의해 볼까:

$$C \in \text{Cmd} \rightarrow \text{Memory} \rightarrow \text{Memory}$$

$$\mathcal{V} \in \text{Expr} \rightarrow \text{Memory} \rightarrow \text{Value}$$

$$C \text{ skip } m = m$$

$$C x := E m = m\{x \mapsto \mathcal{V} E m\}$$

$$C \text{ if } E C_1 C_2 m = \mathcal{V} E m ? C C_1 m : C C_2 m$$

$$C C_1 ; C_2 m = C C_2 (C C_1 m)$$

$$C \text{ while } E \text{ do } C m = \mathcal{V} E m ? C \text{ while } E \text{ do } C (C C m) : m$$

$$\mathcal{V} n m = \{n\}$$

$$\mathcal{V} x m = m x$$

문제

프로그램 C 의 의미는 CC 로 정의되는가? No.

while-문의 경우:

$$\begin{aligned}
 C \text{ while } E \text{ do } C &= \dots C \text{ while } E \text{ do } C \dots \\
 &= \dots (\dots C \text{ while } E \text{ do } C \dots) \dots \\
 &= \dots
 \end{aligned}$$

답) CC 는 명령문 프로그램 C 를 가지고 만드는 방정식을 표현하는 것 뿐. 그 방정식의 해가 프로그램 C 의 의미.

프로그램 C 의 의미 방정식

$$\mathcal{C} C = \dots$$

의 오른쪽은 위에서 표현한 함수 C 의 내용을 고스란히 가지는
상위의 함수

$\mathcal{F} : (Cmd \rightarrow Memory \rightarrow Memory) \rightarrow (Cmd \rightarrow Memory \rightarrow Memory)$

가 정의하는 $\mathcal{F} C C$ 가 된다:

$$\mathcal{F} C \text{ skip } m = m$$

$$\mathcal{F} C x := E m = m\{x \mapsto \mathcal{V} E m\}$$

$$\mathcal{F} C \text{ if } E C_1 C_2 m = \mathcal{V} E m ? \mathcal{C} C_1 m : \mathcal{C} C_2 m$$

$$\mathcal{F} C C_1 ; C_2 m = \mathcal{C} C_2 (\mathcal{C} C_1 m)$$

$$\mathcal{F} C \text{ while } E \text{ do } C m = \mathcal{V} E m ? \mathcal{C} \text{ while } E \text{ do } C (\mathcal{C} C m) : m$$

C 의 의미 방정식

프로그램 C 의 의미는 그 의미 $\mathcal{C}C$ 에 대한 방정식

$$\mathcal{C}C = \mathcal{F}C C$$

와 C 안의 모든 명령문 C_i 들에 대한 방정식

$$\mathcal{C}C_i = \mathcal{F}C C_i$$

들의 해를 가지고 정의된다.

그러한 \mathcal{F} 를 통해서 프로그램 C 로 부터 도출되는 연립방정식을 하나의 함수 \mathcal{F}_C 를 가지고 표현하면

$$\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix} = \mathcal{F}_C \begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}$$

이 되고, 방정식의 주인공 X_i 는 $C C_i$ 를 대신에 쓴 것. (C 안의 명령문들의 갯수는 n)

- \mathcal{F}_C 의 정의는 \mathcal{F} 와 \mathcal{V} 로 부터 명백
- 방정식의 해는 \mathcal{F}_C 의 최소고정점
- \mathcal{F}_C 의 최소고정점은 존재: (why?)

$$\text{lfp} \mathcal{F}_C = \bigsqcup_i (\mathcal{F}_C^i \langle \perp, \dots, \perp \rangle)$$

$lfp\mathcal{F}_C$ 는 $lfp\mathcal{F}$ 의 일부

방정식의 해

$$lfp\mathcal{F}_C$$

는

$$lfp\mathcal{F} \in \text{Cmd} \rightarrow \text{Memory} \rightarrow \text{Memory}$$

중에서 프로그램 C 를 구성하는 명령문들의 의미들로 구성된다:

$$lfp\mathcal{F}_C = \langle (lfp\mathcal{F})C_0, (lfp\mathcal{F})C_1, \dots, (lfp\mathcal{F})C_n \rangle$$

- \mathcal{F}_C 를 구성하는 속 내용은 모두 \mathcal{F} 의 정의 그대로.
- $lfp\mathcal{F} \in \text{Cmd} \rightarrow \text{Memory} \rightarrow \text{Memory}$ 는 모든 명령문에 대한 의미.

앞으로

- 연속인(continuous) 요약 의미 함수를 정의:

$$\hat{\mathcal{F}} \in (Cmd \rightarrow Memory \rightarrow Memory) \rightarrow (Cmd \rightarrow Memory \rightarrow Memory)$$

- 프로그램 C 의 요약 의미는 $\hat{\mathcal{F}}$ 로 부터 도출되는 연립방정식

$$\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix} = \hat{\mathcal{F}}_C \begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}$$

의 해, 즉 $lfp \hat{\mathcal{F}}_C$.

- 방정식의 해 $lfp \hat{\mathcal{F}}_C$ 는 $lfp \mathcal{F}$ 의 일부:

$$lfp \hat{\mathcal{F}}_C = \langle (lfp \hat{\mathcal{F}}) C_0, (lfp \hat{\mathcal{F}}) C_1, \dots, (lfp \hat{\mathcal{F}}) C_n \rangle$$

- 올바른가 검증:

$$\alpha(lfp \mathcal{F}) \sqsubseteq lfp \hat{\mathcal{F}}?$$

즉, 요약해석의 틀에 의해서 다음을 증명하면 됨:

$$\alpha \circ \mathcal{F} \sqsubseteq \hat{\mathcal{F}} \circ \alpha \quad \text{혹은} \quad \alpha(f) \sqsubseteq g \implies \mathcal{F} f \sqsubseteq \hat{\mathcal{F}} g$$

요약된 의미 방정식

프로그램 C 의 요약된 의미는 다음의 요약공간

$$Memory = Loc \xrightarrow{\text{fin}} Value$$

$$Value = \hat{Z} + \hat{B}$$

$$\hat{B} = \{\perp, T, F, \top\}$$

$$Loc = Var$$

을 가지고 다음의 재귀함수의 내용으로 만들어지는 방정식의 해가 되겠다:

$$\hat{C} \in Cmd \rightarrow Memory \rightarrow Memory$$

$$\hat{V} \in Expr \rightarrow Memory \rightarrow Value$$

$$\hat{C} \text{ skip } \hat{m} = \hat{m}$$

$$\hat{C} x := E \hat{m} = \hat{m}\{x \mapsto \hat{V} E \hat{m}\}$$

$$\hat{C} \text{ if } E C_1 C_2 \hat{m} = (\hat{C} C_1 \hat{m}) \sqcup (\hat{C} C_2 \hat{m})$$

$$\hat{C} C_1 ; C_2 \hat{m} = \hat{C} C_2(\hat{C} C_1 \hat{m})$$

$$\hat{C} \text{ while } E \text{ do } C \hat{m} = \hat{m} \sqcup (\hat{C} \text{ while } E \text{ do } C (\hat{C} C \hat{m}))$$

$$\hat{V} n \hat{m} = \alpha\{n\}$$

$$\hat{V} x \hat{m} = \hat{m} x$$

$$\hat{V} E_1 + E_2 \hat{m} = (\hat{V} E_1 \hat{m}) \hat{+} (\hat{V} E_2 \hat{m})$$

$$\hat{V} E_1 < E_2 \hat{m} = (\hat{V} E_1 \hat{m}) \hat{<} (\hat{V} E_2 \hat{m})$$

명령문 C 의 요약해석 방정식

$$\hat{C} C = \dots$$

의 오른쪽은 위에서 표현한 재귀함수 \hat{C} 의 내용을 고스란히 가지는 상위의 함수

$$\hat{\mathcal{F}} : (Cmd \rightarrow Memory \rightarrow Memory) \rightarrow (Cmd \rightarrow Memory \rightarrow Memory)$$

가 정의하는 $\hat{\mathcal{F}} \hat{C} C$ 가 되겠다:

$$\hat{\mathcal{F}} \hat{C} \text{ skip } \hat{m} = \hat{m}$$

$$\hat{\mathcal{F}} \hat{C} x := E \hat{m} = \hat{m} \{x \mapsto \hat{V} E \hat{m}\}$$

$$\hat{\mathcal{F}} \hat{C} \text{ if } E C_1 C_2 \hat{m} = (\hat{C} C_1 \hat{m}) \sqcup (\hat{C} C_2 \hat{m})$$

$$\hat{\mathcal{F}} \hat{C} C_1 ; C_2 \hat{m} = \hat{C} C_2 (\hat{C} C_1 \hat{m})$$

$$\hat{\mathcal{F}} \hat{C} \text{ while } E \text{ do } C \hat{m} = \hat{m} \sqcup (\hat{C} \text{ while } E \text{ do } C (\hat{C} C \hat{m}))$$

요약해석 틀에 의해, $lfp\hat{\mathcal{F}}$ 이 $lfp\mathcal{F}$ 의 안전한 요약이라면, 증명할 것은 아래 둘 중 하나이다:

- $\alpha \circ \mathcal{F} \sqsubseteq \hat{\mathcal{F}} \circ \alpha$, 즉, 모든 프로그램 C 에 대해서

$$\forall f : (\alpha(\mathcal{F} f)) C \sqsubseteq (\hat{\mathcal{F}}(\alpha f)) C,$$

혹은,

- $\alpha f \sqsubseteq \hat{f}$ 이면 $\alpha(\mathcal{F} f) \sqsubseteq \hat{\mathcal{F}} \hat{f}$, 즉, 모든 프로그램 C 에 대해서

$$(\alpha(\mathcal{F} f)) C \sqsubseteq (\hat{\mathcal{F}} \hat{f}) C.$$

여기서

$$(Cmd \rightarrow Memory \rightarrow Memory) \xleftrightarrow[\alpha]{\gamma} (Cmd \rightarrow \hat{Memory} \rightarrow \hat{Memory})$$

이고, 증명은 C 의 각 경우별로 따지면 되는 데, 이 때

$$Memory \xleftrightarrow[\alpha_1]{\gamma_1} \hat{Memory}$$

를 가지고 조립식으로 정의된 α 가 사용된다.

구현

분석의 구현은, 분석할 프로그램 C 가 주어졌을 때 연립방정식

$$\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix} = \hat{\mathcal{F}}_C \begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}$$

을 풀면된다. 연립방정식은 C 안에 있는 모든 명령문들 C_i 와 식들 E_i 의 요약 의미

$$\hat{C} C_i \in \hat{Memory} \rightarrow \hat{Memory}$$

$$\hat{V} E_i \in \hat{Memory} \rightarrow \hat{Value}$$

에 대한 방정식.

방정식의 해(프로그램의 요약된 의미)는 메모리에서 메모리로 가는 함수가 된다.

예

다음 프로그램

$$x := 1 ; \text{while } (0 < x) \text{ do } x := x + 1$$

을 생각하자. 각 부품마다 번호를 붙이자.

$$\underbrace{x := 1}_{1} ; \text{while } (0 < x) \text{ do } \underbrace{x := x + 1}_{4}$$

$$\underbrace{\hspace{10em}}_{3}$$

$$\underbrace{\hspace{15em}}_{2}$$

$$\underbrace{\hspace{20em}}_{0}$$

각 부품의 의미 $\hat{C} C_i$ 와 $\hat{V} E_i$ 에 대한 방정식은 각각 $\hat{F} C_i$ 와 $\hat{V} E_i$ 에 의해서 아래와 같이 정의된다:

$$\begin{aligned} \hat{C} C_0 &= \lambda \hat{m}. \hat{C} C_2 (\hat{C} C_1 \hat{m}) \\ \hat{C} C_1 &= \lambda \hat{m}. \hat{m} \{x \mapsto \hat{V} 1 \hat{m}\} \\ &= \lambda \hat{m}. \hat{m} \{x \mapsto \alpha\{1\}\} \\ \hat{C} C_2 &= \lambda \hat{m}. \hat{m} \sqcup (\hat{C} C_2 (\hat{C} C_3 \hat{m})) \\ \hat{C} C_3 &= \lambda \hat{m}. \hat{m} \{x \mapsto \hat{V} E_4 \hat{m}\} \\ &= \lambda \hat{m}. \hat{m} \{x \mapsto (\hat{m} x) \hat{\dagger} \alpha\{1\}\} \end{aligned}$$

방정식의 주인공 $\hat{C} C_i$ 를 \hat{X}_i 로 해서 다시 쓰면,

$$\begin{aligned} \hat{X}_0 &= \lambda \hat{m}. \hat{X}_2 (\hat{X}_1 \hat{m}) \\ \hat{X}_1 &= \lambda \hat{m}. \hat{m} \{x \mapsto \alpha\{1\}\} \\ \hat{X}_2 &= \lambda \hat{m}. \hat{m} \sqcup (\hat{X}_2 (\hat{X}_3 \hat{m})) \\ \hat{X}_3 &= \lambda \hat{m}. \hat{m} \{x \mapsto (\hat{m} x) \hat{\dagger} \alpha\{1\}\} \end{aligned}$$

$2^{\mathbb{Z}} \begin{matrix} \xleftarrow{\gamma} \\ \xrightarrow{\alpha} \end{matrix} \hat{\mathbb{Z}}$ 의 두가지 경우를 살펴보자

- $\hat{\mathbb{Z}}$ 의 높이가 유한한 경우
- $\hat{\mathbb{Z}}$ 의 높이가 무한한 경우

$$\hat{X}_0 = \lambda \hat{m}. \hat{X}_2(\hat{X}_1 \hat{m})$$

$$\hat{X}_1 = \lambda \hat{m}. \hat{m}\{x \mapsto \alpha\{1\}\}$$

$$\hat{X}_2 = \lambda \hat{m}. \hat{m} \sqcup (\hat{X}_2 (\hat{X}_3 \hat{m}))$$

$$\hat{X}_3 = \lambda \hat{m}. \hat{m}\{x \mapsto (\hat{m} x) \hat{\dagger} \alpha\{1\}\}$$

\hat{Z} 의 높이가 유한한 경우

예를들어, $\hat{Z} = \{\perp, -, +, \top\}$ 일 때, 위의 방정식은:

$$\hat{X}_0 = \lambda \hat{m}. \hat{X}_2(\hat{X}_1 \hat{m})$$

$$\hat{X}_1 = \lambda \hat{m}. \hat{m}\{x \mapsto +\}$$

$$\hat{X}_2 = \lambda \hat{m}. \hat{m} \sqcup (\hat{X}_2 (\hat{X}_3 \hat{m}))$$

$$\hat{X}_3 = \lambda \hat{m}. \hat{m}\{x \mapsto (\hat{m} x) \hat{+} +\}$$

프로그램 시작에서의 모든 메모리 상태를 포섭하는 것이

$$\{\} \in Memory = Var \xrightarrow{\text{fin}} Value$$

라고하면,

$$\hat{X}_0 \{\}$$

에서 부터 “연쇄반응”을 일으키는 방정식들만 풀다.

연쇄반응을 따라, 풀어야 할 방정식들:

$$\hat{X}_0 \{\} = \hat{X}_2(\hat{X}_1 \{\}) \quad \text{따라서}$$

$$\hat{X}_1 \{\} = \{x \mapsto +\} \quad \text{따라서}$$

$$\hat{X}_2 \{x \mapsto +\} = \{x \mapsto +\} \sqcup (\hat{X}_2(\hat{X}_3 \{x \mapsto +\})) \quad \text{따라서}$$

$$\hat{X}_3 \{x \mapsto +\} = \{x \mapsto +\} \{x \mapsto + \hat{\dagger} +\}$$

다시 정리하면

$$\hat{X}_0 \{\} = \hat{X}_2 \{x \mapsto +\}$$

$$\hat{X}_1 \{\} = \{x \mapsto +\}$$

$$\hat{X}_2 \{x \mapsto +\} = \{x \mapsto +\} \sqcup (\hat{X}_2 \{x \mapsto +\})$$

$$\hat{X}_3 \{x \mapsto +\} = \{x \mapsto +\}$$

위의 방정식을 다시 쓰면 (\hat{X}_i ★을 \hat{Y}_i 로)

$$\hat{Y}_0 = \hat{Y}_2$$

$$\hat{Y}_1 = \{x \mapsto +\}$$

$$\hat{Y}_2 = \{x \mapsto +\} \sqcup \hat{Y}_2$$

$$\hat{Y}_3 = \{x \mapsto +\}$$

위의 방정식의 최소해는 고정점 계산(*fixpoint iteration*)

($\sqcup_i (f^i \hat{1})$ 의 계산) 으로:

$$\hat{Y}_0 = \{x \mapsto +\}$$

$$\hat{Y}_1 = \{x \mapsto +\}$$

$$\hat{Y}_2 = \{x \mapsto +\}$$

$$\hat{Y}_3 = \{x \mapsto +\}$$

분석결과: 프로그램의 모든 명령문 실행 후 x 는 음이아닌 정수를 가진다.