Week 5.2

Thursday, February 8, 2024 4:58 PM

DFA limitations so far:

- pointers
- function calls

DFA pointer analysis

- L possible
- L costly

inclusion-based set constraint analysis

DFA = flow-sensitive

set constraint = flow-insensitive

language

$x \in$ Variable

$c \in$ Constructor

$t \in \text{Term} ::= x \mid c(t_1, \ldots, t_n)$ ← constructor call

$e \in \text{Exp} ::= t \mid c^{-i}(x)$ ← projection

$s \in \text{Stmt} ::= e_1 \subseteq e_2 \mid s_1 \wedge s_2$

e_1 & e_2 cannot both

e_1 & e_2 cannot both be projections

---

constructor = uninterpreted function

ex type constructors

nullary constructors : int, bool, string

unary constructors : Set[·], List[·]

binary constructors : · → ·, Map[·,·]

2 characteristics:

- arity : # arguments
- variance : each argument position is either
  - covariant : monotone wrt that position
  - contravariant : anti-tone wrt that position

variance is answering the following question:

"if we know $A \subseteq B$, what is true about the relation between $c(A)$ & $c(B)$?" → (or vice-versa)

EXAMPLE

arity = 2

Consider foo/2 where foo's 1st position is covariant & its 2nd position is contravariant

suppose $foo(A, B) \subseteq foo(C, D)$

↳ $A \subseteq C$ (because covariant)

↳ $A \subseteq C$ (because covariant)

↳ $D \subseteq B$ (because contravariant)

projections

$c^{-i}(X)$: take value of X (which is a set of constructor calls), filter them to keep only 'c' constructor calls, take the i'th argument of each such call & put them all in a set

example:

$X \mapsto \{a(b,c), a(D,e), f(g), h(i,j,k)\}$

$C \mapsto \{l, m\}$

$D \mapsto \{n, o, p\}$

then $a^{-1}(X) = \{C, e\} = \{e, l, m\}$

---

example

$f(X,Y) \subseteq A$ | $h \subseteq f^{-1}(B)$

$A \subseteq B$ | $D \subseteq f^{-0}(C)$

$A \subseteq B$
$B \subseteq C$
$C \subseteq E$
$g \subseteq D$
$h \subseteq D$

$D \subseteq +^{-}(C)$
$f^{-\phi}(C) \subseteq E$
$f^{-1}(B) \subseteq F$

what is a solution?

- let H be the set of constructor calls
- a solution is an assignment $\sigma \in \text{Variable} \to \mathcal{P}(H)$ s.t. all constraints are satisfied

$A \mapsto \{f(X,Y)\}$
$B \mapsto \{f(X,Y)\}$
$C \mapsto \{f(X,Y)\}$
$D \mapsto \{g, h\}$

$E \mapsto \{f(X,Y), g, h\}$
$F \mapsto \{h\}$
$X \mapsto \{g, h\}$
$Y \mapsto \{h\}$

---

this language is:

- decidable
- guaranteed to have a minimal soln
- $O(n^3)$

analyses in this language are called "inclusion-based analyses"

↳ sidenote: can replace "$e_1 \subseteq e_2$" w/ "$e_1 = e_2$", get equality-based analysis

↳ lost precision

↳ near-linear complexity

$O(n \cdot \alpha(n))$

Solving Constraints

- convenient to represent constraints as a graph
  - ↳ node for each variable, call, & projection
  - ↳ edges for constraints
- kinds of edges to represent $a \to b$
  - ↳ successor edges: store edge info in a
  - ↳ predecessor edges: store edge info in b
- edge rules:
  - . any edge from call is pred. edge
  - . any edge to a proj is a pred. edge

- any edge to a proj is a pred. edge
- all other edges are succ. edges

ex