CS-XXX: Graduate Programming Languages

Lecture 5 — Pseudo-Denotational Semantics

Dan Grossman
2012
A different approach

Operational semantics defines an interpreter, from abstract syntax to abstract syntax. Metalanguage is inference rules (slides) or OCaml (interp.ml)

Denotational semantics defines a compiler (translater), from abstract syntax to a different language with known semantics

Target language is math, but we’ll make it a tiny core of OCaml (hence “pseudo”)

Metalanguage is math or OCaml (we’ll show both)
The basic idea

A heap is a math/ML function from strings to integers:

\[ \text{string} \rightarrow \text{int} \]

An expression denotes a math/ML function from heaps to integers

\[ \text{den}(e) : (\text{string} \rightarrow \text{int}) \rightarrow \text{int} \]

A statement denotes a math/ML function from heaps to heaps

\[ \text{den}(s) : (\text{string} \rightarrow \text{int}) \rightarrow (\text{string} \rightarrow \text{int}) \]

Now just define \textit{den} in our metalanguage (math or ML), inductively over the source language abstract syntax
Expressions

\[ den(e) : (\text{string} \rightarrow \text{int}) \rightarrow \text{int} \]

\[
\begin{align*}
\text{den}(c) &= \text{fun } h \rightarrow c \\
\text{den}(x) &= \text{fun } h \rightarrow h\ x \\
\text{den}(e_1 + e_2) &= \text{fun } h \rightarrow (\text{den}(e_1)\ h) + (\text{den}(e_2)\ h) \\
\text{den}(e_1 \ast e_2) &= \text{fun } h \rightarrow (\text{den}(e_1)\ h) \ast (\text{den}(e_2)\ h)
\end{align*}
\]

In plus (and times) case, two “ambiguities”:

- “+” from meta language or target language?
  - Translate abstract + to OCaml +, (ignoring overflow)

- When do we denote \(e_1\) and \(e_2\)?
  - Not a focus of the metalanguage. At “compile time”.

Dan Grossman
CS-XXX 2012, Lecture 5
Switching metalanguage

With OCaml as our metalanguage, ambiguities go away

But it is harder to distinguish mentally between “target” and “meta”

If denote in function body, then source is “around at run time”

▶ After translation, should be able to “remove” the definition of the abstract syntax

▶ ML does not have such a feature, but the point is we no longer need the abstract syntax

See denote.ml
Statements, w/o while

\[ den(s) : (\text{string} \rightarrow \text{int}) \rightarrow (\text{string} \rightarrow \text{int}) \]

\begin{align*}
  den(\text{skip}) & = \text{fun } h \rightarrow h \\
  den(x := e) & = \\
  & \quad \text{fun } h \rightarrow (\text{fun } v \rightarrow \text{if } x = v \text{ then } den(e) h \text{ else } h v) \\
  den(s_1; s_2) & = \text{fun } h \rightarrow den(s_2) (den(s_1) h) \\
  den(\text{if } e s_1 s_2) & = \\
  & \quad \text{fun } h \rightarrow \text{if } den(e) h > 0 \text{ then } den(s_1) h \text{ else } den(s_2) h \\
\end{align*}

Same ambiguities; same answers

See denote.ml
While

\[ \text{den(while } e \ s) = \ |
\]
\[
\begin{align*}
\text{let rec } & f \ h = & \text{let d1=denote_exp } e \text{ in} \\
\text{if } (\text{den}(e) \ h) > 0 & \text{ then } f (\text{den}(s) \ h) & \text{let d2=denote_stmt } s \text{ in} \\
\text{else } h \text{ in } & f & \text{let rec } f \ h = \\
\end{align*}
\]
\[
\begin{align*}
\text{if } (d1 \ h) > 0 & \text{ then } f (d2 \ h) \\
\text{else } h \text{ in } & f \\
\end{align*}
\]

The function denoting a while statement is inherently recursive!

Good thing our target language has recursive functions!

Why doesn’t \[ \text{den(while } e \ s) = \text{den(if } e (s; \text{while } e \ s) \text{ skip)} \] make any sense?
Two common mistakes

A denotational semantics should “eagerly” translate the entire program
  ▶ E.g., both branches of an if

But a denotational semantics should “terminate”
  ▶ I.e., avoid any circular definitions in the translating
  ▶ The result of the translation can use (well-founded) recursion
  ▶ E.g., compiling a while-loop should not produce an infinite amount of code
Finishing the story

```ocaml
let denote_prog s = 
    let d = denote_stmt s in 
    fun () -> (d (fun x -> 0)) "ans"
```

Compile-time: `let x = denote_prog (parse file)`

Run-time: `print_int (x ())`

In-between: We have an OCaml program using only functions, variables, ifs, constants, +, *, >, etc.

- Does not use any constructors of exp or stmt (e.g., Seq)
The real story

For “real” denotational semantics, target language is math

(And we write $[s]$ instead of $\text{den}(s)$)

Example: $[x := e][H] = [H][x \mapsto [e][H]]$

There are two major problems, both due to while:

1. Math functions do not diverge, so no function denotes \textbf{while 1 skip}
2. The denotation of loops cannot be circular
The elevator version, which we will not pursue

For (1), we “lift” the semantic domains to include a special $\bot$

$\text{den}(s) : (\text{string} \rightarrow \text{int}) \rightarrow (((\text{string} \rightarrow \text{int}) \cup \bot)$

- Have to change meaning of $\text{den}(s_2) \circ \text{den}(s_1)$ appropriately

For (2), we use $\textbf{while } e \ s$ to define a (meta)function $f$ that given a lifted heap-transformer $X$ produces a lifted heap-transformer $X'$:

- If $\text{den}(e)(\text{den}(H)) = 0$, then $\text{den}(H)$
- Else $X \circ \text{den}(s)$

Now let $\text{den}(\textbf{while } e \ s)$ be the least fixed-point of $f$

- An hour of math to prove the least fixed-point exists
- Another hour to prove it is the limit of starting with $\bot$ and applying $f$ over and over (i.e., any number of loop iterations)
- Keywords: monotonic functions, complete partial orders, Knaster-Tarski theorem
Where we are

- Have seen operational and denotational semantics
- Connection to interpreters and compilers
- Useful for rigorous definitions and proving properties
- Next: Equivalence of semantics
  - Crucial for compiler writers
  - Crucial for code maintainers
- Then: Leave IMP behind and consider functions

But first: Will any of this help write an O/S service?