Comp 411
Principles of Programming Languages
Lecture 25
The CPS Transformation
(Reprise of Lecture 24)

Corky Cartwright
April 3, 2020
CPS Granularity

In pure form, the CPS transformation is typically given for the untyped $\lambda$-calculus (see the optional notes on the CPS Transformation in OCaml). But this characterization (like most formalisms based on the untyped $\lambda$-calculus) is misleading in practice because it does not address the issue of processing primitive operations (the untyped $\lambda$-calculus has no primitive operations!). Neither does the polymorphic $\lambda$-calculus (System F).

Of course, primitive operations are much easier to process than program functions because they typically do not abort (a few operations like division and object accessors are exceptions) or otherwise discard the pending continuation. Modular 2’s complement arithmetic (other than division) is a good example.

But primitive operations can be treated like program functions provided the libraries implementing are re-shaped so that every such operation takes an extra continuation argument. The designation of which operations are primitive has a huge impact on the final form of the CPSed code. If primitive operations are CPSed, then the CPSed code is much more complex. In practice, CPSing primitives is generally not advisable since CPSing adds overhead (extra function arguments and extra function calls) and we typically only need to CPS the operations that correspond to subroutine calls.
The CPS transformation is often performed by compilers for “higher-order” languages (those that support functions as data values), because CPSing exposes all of the operations that are implicitly performed on the stack in standard code (which uses an algol-like stack run-time).

But there are less severe alternative transformations (notably A-normal form) that perform much the same function. In A-normal form, every non-trivial intermediate result is explicitly stored in a local variable. An application is trivial iff the rator is a primitive operation.

If no operation is treated as primitive, then A-normal form conversion is very similar to a much older representation used in optimizing compilers called value-numbering. In value-numbering, hashing is used to avoid duplicating subtrees in a concrete representation of the abstract syntax of a program.
Review: The CPS Transformation

Assume Jam/Scheme programs are restricted to a form where the body of a function is either (i) a primitive expression constructed from constants, variables and primitive functions, and program-defined functions; or (ii) a conditional where the predicates are *primitive* expressions and the result clauses are *ordinary* expressions (primitive expressions augmented by program-defined functions). Then the CPS transformation of such a program is defined as follows:

1. Add an extra parameter $k$ to every function.

2. For each function body $b$ that is a primitive expression, write $k(b)$.

3. Each clause in a conditional is treated separately:
   
   a. For each result clause $b$ composed from primitive operations and constants, write $k(b)$.

   b. For each clause containing calls on program-defined functions, pick the call that will be evaluated first. Make the body of the new clause a call on a reshaped version of the program-defined function that takes an extra argument of the form $\text{map res to body}$, called the continuation. The original contents of that clause are placed in the body, enclosed in a call on the continuation $k$, with the selected call replaced by $\text{res}$.

   c. Repeat preceding step 3b) until no unconverted function calls remain.
Review: Another Example

```
let treeSum :=
    map t to if leaf?(t) then t
    else Tree-Sum(left(t)) + Tree-Sum(right(t))
In treeSum( ... )
```

Then first iteration in creating the CPS version, \texttt{treeSumK}, is

```
let treeSumK :=
    map t,k to if leaf?(t) then k(t)
    else treeSumK(left(t),
               map res to k(res + treeSum(right(t))))
in treeSumK( ... , map x to x)
```
let

    treeSumK := map t,k to  // rule 1
        if leaf?(t) k(t)       // rule 3a
        else treeSumK(left(t),  // rule 3b
            map r1 to treeSumK(right(t),
                map r2 to k(r1 + r2)))

in treeSumK(..., map x to x)
Comprehensive Formulations of the CPS Transformation

The rules for performing the CPS transformation are more complex in the context of explicit binding constructs like \texttt{lambda}, \texttt{let}, and \texttt{letrec} (recursive \texttt{let}). In principle, these extensions do not add anything new, but they complicate the detailed structure of environments and the CPS transformations eliminates explicit environments (other than local variables) by encoding environments (represented using the stack in algol-like run-times) as closures (continuations) in the heap.

Study the rules for Assignment 6, which constitute one possible way to handle the Jam recursive \texttt{let} and \texttt{map} constructs. Good CPS translations are concise. The rules for Assignment 6 produce reasonably concise CPS translations but they could be improved at the cost of more complexity.