CS 61A Week 1 First Lab

Monday afternoon, Tuesday, or Wednesday morning

Try to get as much done as possible, but don't panic if you don't finish everything.

- 1. Start the Emacs editor, either by typing emacs in your main window or by selecting it from the alt-middle mouse menu. (Your TA will show you how to do this.) From the Help menu, select the Emacs tutorial. You need not complete the entire tutorial at the first session, but you should do so eventually.
- 2. Start Scheme, either by typing stk in your main window or by typing meta-S in your Emacs window. Type each of the following expressions into Scheme, ending the line with the Enter (carriage return) key. Think about the results! Try to understand how Scheme interprets what you type.

```
(first 'hello)
(+23)
                                    (first hello)
(+5678)
                                    (first (bf 'hello))
(+)
                                    (+ (first 23) (last 45))
(sqrt 16)
                                    (define pi 3.14159)
(+ (* 3 4) 5)
                                    'pi
, +
                                    (+ pi 7)
'hello
                                    (* pi pi)
'(+23)
                                    (define (square x) (* x x))
'(good morning)
                                    (square 5)
(first 274)
                                    (square (+ 2 3))
(butfirst 274)
```

3. Use Emacs to create a file called pigl.scm in your directory containing the Pig Latin program shown below:

```
(define (pigl wd)
  (if (pl-done? wd)
        (word wd 'ay)
        (pigl (word (bf wd) (first wd)))))
(define (pl-done? wd)
  (vowel? (first wd)))
(define (vowel? letter)
  (member? letter '(a e i o u)))
```

Make sure you are editing a file whose name ends in .scm, so that Emacs will know to indent your code correctly!

4. Now run Scheme. You are going to create a transcript of a session using the file you just created:

5. Use lpr to print your transcript file.

CS 61A Week 1 Second Lab Wednesday afternoon, Thursday, or Friday morning

1. Predict what Scheme will print in response to each of these expressions. *Then* try it and make sure your answer was correct, or if not, that you understand why!

```
(define a 3)
(define b (+ a 1))
(+ a b (* a b))
(= a b)
(if (and (> b a) (< b (* a b)))
    b
    a)
(cond ((= a 4) 6)
        ((= b 4) (+ 6 7 a))
        (else 25))
(+ 2 (if (> b a) b a))
        ((< a b) b)
              (else -1))
        (+ a 1))
((if (< a b) + -) a b)</pre>
```

2. In the shell, type the command

```
cp ~cs61a/lib/plural.scm .
```

(Note the period at the end of the line!) This will copy a file from the class library to your own directory. Then, using emacs to edit the file, modify the procedure so that it correctly handles cases like (plural 'boy).

- **3.** Define a procedure that takes three numbers as arguments and returns the sum of the squares of the two larger numbers.
- **4.** Write a procedure dupls-removed that, given a sentence as input, returns the result of removing duplicate words from the sentence. It should work this way:

```
> (dupls-removed '(a b c a e d e b))
(c a d e b)
> (dupls-removed '(a b c))
(a b c)
> (dupls-removed '(a a a a b a a))
(b a)
```

CS 61A Week 2 Lab Monday afternoon, Tuesday, or Wednesday morning

This lab introduces a new special form, lambda.

1. Type each of the following into Scheme, and note the results. See when you can predict the results before letting Scheme do the computation.

```
(lambda (x) (+ x 3))

((lambda (x) (+ x 3) 7)

You can think of lambda as meaning "the function of...," e.g., "the function of x that returns (+ x 3).

(define (make-adder num)
    (lambda (x) (+ x num)))

((make-adder 3) 7)

(define plus3 (make-adder 3))

(plus3 7)

(define (square x) (* x x))

(square 5)

(define sq (lambda (x) (* x x)))

(sq 5)

(define (try f) (f 3 5))

(try +)
```

Week 2 continued...

- 2. Write a procedure **substitute** that takes three arguments: a sentence, an *old* word, and a *new* word. It should return a copy of the sentence, but with every occurrence of the old word replaced by the new word. For example:
- > (substitute '(she loves you yeah yeah yeah) 'yeah 'maybe)
 (she loves you maybe maybe maybe)
- 3. Consider a Scheme function g for which the expression

```
((g) 1)
```

returns the value 3 when evaluated. Determine how many arguments g has. In one word, also describe as best you can the *type* of value returned by g.

4. For each of the following expressions, what must f be in order for the evaluation of the expression to succeed, without causing an error? For each expression, give a definition of f such that evaluating the expression will not cause an error, and say what the expression's value will be, given your definition.

```
f
(f)
(f 3)
((f))
(((f)) 3)
```

5. Find the values of the expressions

```
((t 1+) 0) ((t (t 1+)) 0) (((t t) 1+) 0)
```

where 1+ is a primitive procedure that adds 1 to its argument, and t is defined as follows:

```
(define (t f)
  (lambda (x) (f (f (f x)))) )
```

Work this out yourself before you try it on the computer!

6. Find the values of the expressions

```
((t s) 0) ((t (t s)) 0) (((t t) s) 0)
```

where t is defined as in question 2 above, and s is defined as follows:

```
(define (s x)
(+ 1 x))
```

7. Write and test the make-tester procedure. Given a word w as argument, make-tester returns a procedure of one argument x that returns true if x is equal to w and false otherwise. Examples:

```
> ((make-tester 'hal) 'hal)
#t
> ((make-tester 'hal) 'cs61a)
#f
> (define sicp-author-and-astronomer? (make-tester 'gerry))
> (sicp-author-and-astronomer? 'hal)
#f
> (sicp-author-and-astronomer? 'gerry)
#t
```

CS 61A Week 3 Lab Monday afternoon, Tuesday, or Wednesday morning

This lab exercise concerns the change counting program on pages 40–41 of Abelson and Sussman.

- 1. Identify two ways to change the program to *reverse* the order in which coins are tried, that is, to change the program so that pennies are tried first, then nickels, then dimes, and so on.
- 2. Abelson and Sussman claim that this change would not affect the *correctness* of the computation. However, it does affect the *efficiency* of the computation. Implement one of the ways you devised in exercise 1 for reversing the order in which coins are tried, and determine the extent to which the number of calls to cc is affected by the revision. Verify your answer on the computer, and provide an explanation. Hint: limit yourself to nickels and pennies, and compare the trees resulting from (cc 5 2) for each order.
- 3. Modify the cc procedure so that its kinds-of-coins parameter, instead of being an integer, is a sentence that contains the values of the coins to be used in making change. The coins should be tried in the sequence they appear in the sentence. For the count-change procedure to work the same in the revised program as in the original, it should call cc as follows:

```
(define (count-change amount)
  (cc amount '(50 25 10 5 1)) )
```

4. Many Scheme procedures require a certain type of argument. For example, the arithmetic procedures only work if given numeric arguments. If given a non-number, an error results.

Suppose we want to write *safe* versions of procedures, that can check if the argument is okay, and either call the underlying procedure or return **#f** for a bad argument instead of giving an error. (We'll restrict our attention to procedures that take a single argument.)

```
> (sqrt 'hello)
ERROR: magnitude: Wrong type in arg1 hello
> (type-check sqrt number? 'hello)
#f
> (type-check sqrt number? 4)
2
```

Write type-check. Its arguments are a function, a type-checking predicate that returns #t if and only if the datum is a legal argument to the function, and the datum.

5. We really don't want to have to use type-check explicitly every time. Instead, we'd like to be able to use a safe-sqrt procedure:

```
> (safe-sqrt 'hello)
#f
> (safe-sqrt 4)
2
```

Don't write safe-sqrt! Instead, write a procedure make-safe that you can use this way:

```
> (define safe-sqrt (make-safe sqrt number?))
```

It should take two arguments, a function and a type-checking predicate, and return a new function that returns **#f** if its argument doesn't satisfy the predicate.

CS 61A Week 4 Lab Monday afternoon, Tuesday, or Wednesday morning

1. Try these in Scheme:

```
(define x (cons 4 5))
(car x)
(cdr x)
(define y (cons 'hello 'goodbye))
(define z (cons x y))
(car (cdr z))
(cdr (cdr z))
2. Predict the result of each of these before you try it:
(cdr (car z))
(car (cons 8 3))
(car z)
(car 3)
3. Enter these definitions into Scheme:
(define (make-rational num den)
  (cons num den))
(define (numerator rat)
  (car rat))
(define (denominator rat)
  (cdr rat))
(define (*rat a b)
  (make-rational (* (numerator a) (numerator b))
                  (* (denominator a) (denominator b))))
(define (print-rat rat)
  (word (numerator rat) '/ (denominator rat)))
4. Try this:
(print-rat (make-rational 2 3))
(print-rat (*rat (make-rational 2 3) (make-rational 1 4)))
```

- 5. Define a procedure +rat to add two rational numbers, in the same style as *rat above.
- 6. Now do exercises 2.2, 2.3, and 2.4 from SICP.
- 7. This week you'll learn that sentences are a special case of *lists*, which are built out of pairs. Explore how that's done with experiments such as these:

```
(define x '(a (b c) d))
(car x)
(cdr x)
(car (cdr x))
```

8. SICP ex. 2.18; this should take some thought, and you should make sure you get it right, but don't get stuck on it for the whole hour. **Note:** Your solution should reverse *lists*, not sentences! That is, you should be using cons, car, and cdr, not first, sentence, etc.

CS 61A Week 5 Lab Monday afternoon, Tuesday, or Wednesday morning

- 1. SICP ex. 2.25 and 2.53; these should be quick and easy.
- 2. SICP ex. 2.55; explain your answer to your TA.
- 3. SICP ex. 2.27. This is the central exciting adventure of today's lab! Think hard about it.
- 4. Each person individually make up a procedure named mystery that, given two lists as arguments, returns the result of applying *exactly two* of cons, append, or list to mystery's arguments, using no quoted values or other procedure calls. Here are some examples of what is and is not fair game:

```
okay

(define (mystery L1 L2) (define (mystery L1 L2) (cons L1 (append L2 L1)))

(define (mystery L1 L2) (define (mystery L1 L2) (list L1 (list L1 L1)))

(define (mystery L1 L2) (cons L1 L2))

(define (mystery L1 L2) (define (mystery L1 L2) (define (mystery L1 L2) (append (cons L2 L2) L1))
```

Type your mystery definition into a file, and have one of your partners load it into Scheme and try to guess what it is by trying it out with various arguments.

After everyone has tried someone else's procedure, decide with your partners which procedure was hardest to guess and why, and what test cases were most and least helpful in revealing the definitions.

CS 61A Week 6 Lab

Monday afternoon, Tuesday, or Wednesday morning

1. Load the Scheme-1 interpreter from the file

~cs61a/lib/scheme1.scm

To start the interpreter, type (scheme-1). Familiarize yourself with it by evaluating some expressions. Remember: you have all the Scheme primitives for arithmetic and list manipulation; you have lambda but not higher-order functions; you don't have define. To stop the scheme-1 interpreter and return to STk, just evaluate an illegal expression, such as ().

1a. Trace in detail how a simple procedure call such as

```
((lambda (x) (+ x 3)) 5)
```

is handled in scheme-1.

1b. Try inventing higher-order procedures; since you don't have **define** you'll have to use the Y-combinator trick, like this:

1c. Since all the Scheme primitives are automatically available in **scheme-1**, you might think you could use **STk**'s primitive **map** function. Try these examples:

```
Scheme-1: (map first '(the rain in spain))
Scheme-1: (map (lambda (x) (first x)) '(the rain in spain))
```

Explain the results.

1d. Modify the interpreter to add the **and** special form. Test your work. Be sure that as soon as a false value is computed, your **and** returns **#f** without evaluating any further arguments.

For the rest of the lab, start by reading SICP section 2.3.3 (pages 151–161).

- 2. SICP ex. 2.62.
- 3. The file ~cs61a/lib/bst.scm contains the binary search tree procedures from pages 156-157 of SICP. Using adjoin-set, construct the trees shown on page 156.

CS 61A Week 7 Lab Monday afternoon, Tuesday, or Wednesday morning

1. Modify the person class given in the lecture notes for week 7 (it's in the file demo2.scm in the ~cs61a/lectures/3.0 directory) to add a repeat method, which repeats the last thing said. Here's an example of responses to the repeat message.

```
> (define brian (instantiate person 'brian))
brian
> (ask brian 'repeat)
()
> (ask brian 'say '(hello))
(hello)
> (ask brian 'repeat)
(hello)
> (ask brian 'greet)
(hello my name is brian)
> (ask brian 'repeat)
(hello my name is brian)
> (ask brian 'ask '(close the door))
(would you please close the door)
> (ask brian 'repeat)
(would you please close the door)
```

2. This exercise introduces you to the usual procedure described on page 9 of "Object-oriented Programming – Above-the-line View". Read about usual there to prepare for lab.

Suppose that we want to define a class called double-talker to represent people that always say things twice, for example as in the following dialog.

```
> (define mike (instantiate double-talker 'mike))
mike
> (ask mike 'say '(hello))
(hello hello)
> (ask mike 'say '(the sky is falling))
(the sky is falling the sky is falling)
Consider the following three definitions for the double-talker class. (They can be found online in the file
~cs61a/lib/double-talker.scm.)
(define-class (double-talker name)
  (parent (person name))
  (method (say stuff) (se (usual 'say stuff) (ask self 'repeat))) )
(define-class (double-talker name)
  (parent (person name))
  (method (say stuff) (se stuff stuff)) )
(define-class (double-talker name)
  (parent (person name))
  (method (say stuff) (usual 'say (se stuff stuff))) )
```

Determine which of these definitions work as intended. Determine also for which messages the three versions would respond differently.

CS 61A Week 8 Lab

Monday afternoon, Tuesday, or Wednesday morning

1. Given below is a simplified version of the make-account procedure on page 223 of Abelson and Sussman.

```
(define (make-account balance)
  (define (withdraw amount)
      (set! balance (- balance amount)) balance)
  (define (deposit amount)
      (set! balance (+ balance amount)) balance)
  (define (dispatch msg)
      (cond
            ((eq? msg 'withdraw) withdraw)
            ((eq? msg 'deposit) deposit) ))
  dispatch)
```

Fill in the blank in the following code so that the result works exactly the same as the make-account procedure above, that is, responds to the same messages and produces the same return values. The differences between the two procedures are that the inside of make-account above is enclosed in the let below, and the names of the parameter to make-account are different.

2. Modify either version of make-account so that, given the message balance, it returns the current account balance, and given the message init-balance, it returns the amount with which the account was initially created. For example:

```
> (define acc (make-account 100))
acc
> (acc 'balance)
100
```

Week 8 continued:

3. Modify make-account so that, given the message transactions, it returns a list of all transactions made since the account was opened. For example:

```
> (define acc (make-account 100))
acc
> ((acc 'withdraw) 50)
50
> ((acc 'deposit) 10)
60
> (acc 'transactions)
((withdraw 50) (deposit 10))
4. Given this definition:
(define (plus1 var)
   (set! var (+ var 1))
   var)
Show the result of computing
(plus1 5)
```

using the substitution model. That is, show the expression that results from substituting 5 for var in the body of plus1, and then compute the value of the resulting expression. What is the actual result from Scheme?

Week 8 continued:

This lab activity consists of example programs for you to run in Scheme. Predict the result before you try each example. If you don't understand what Scheme actually does, ask for help! Don't waste your time by just typing this in without paying attention to the results.

```
(define (make-adder n)
                                                ((lambda (x)
  (lambda (x) (+ x n)))
                                                   (let ((a 3))
                                                     (+ x a)))
(make-adder 3)
                                                5)
((make-adder 3) 5)
                                                (define k
                                                  (let ((a 3))
                                                    (lambda (x) (+ x a))))
(define (f x) (make-adder 3))
(f 5)
                                                (k 5)
(define g (make-adder 3))
                                                (define m
                                                  (lambda (x)
(g 5)
                                                    (let ((a 3))
                                                      (+ x a))))
(define (make-funny-adder n)
  (lambda (x)
                                                (m 5)
    (if (equal? x 'new)
        (set! n (+ n 1))
                                                (define p
        (+ x n))))
                                                  (let ((a 3))
                                                    (lambda (x)
(define h (make-funny-adder 3))
                                                      (if (equal? x 'new)
                                                          (set! a (+ a 1))
(define j (make-funny-adder 7))
                                                          (+ x a)))))
(h 5)
                                                (p 5)
(h 5)
                                                (p 5)
(h 'new)
                                                (p 'new)
(h 5)
                                                (p 5)
(j 5)
                                                (define r
                                                  (lambda (x)
(let ((a 3))
                                                    (let ((a 3))
 (+ 5 a))
                                                      (if (equal? x 'new)
                                                          (set! a (+ a 1))
(let ((a 3))
                                                          (+ x a)))))
  (lambda (x) (+ x a)))
                                                (r 5)
((let ((a 3))
   (lambda (x) (+ x a)))
                                                (r 5)
5)
                                                (r 'new)
                                                (r 5)
```

Week 8 continued:

```
(define s
                                                (define (ask obj msg . args)
  (let ((a 3))
                                                  (apply (obj msg) args))
    (lambda (msg)
      (cond ((equal? msg 'new)
                                                (ask s 'add 5)
             (lambda ()
                (set! a (+ a 1))))
                                                (ask s 'new)
            ((equal? msg 'add)
(lambda (x) (+ x a)))
                                                (ask s 'add 5)
             (else (error "huh?"))))))
                                                (define x 5)
(s 'add)
                                                (let ((x 10)
(s 'add 5)
                                                      (f (lambda (y) (+ x y))))
                                                  (f 7))
((s 'add) 5)
                                                (define x 5)
(s 'new)
((s 'add) 5)
((s 'new))
((s 'add) 5)
```

CS 61A Week 9 Lab Monday afternoon, Tuesday, or Wednesday morning

- 1. Exercise 3.12 of Abelson and Sussman.
- 2. Suppose that the following definitions have been provided.

```
(define x (cons 1 3))
(define y 2)
```

A CS 61A student, intending to change the value of x to a pair with car equal to 1 and cdr equal to 2, types the expression (set! (cdr x) y) instead of (set-cdr! x y) and gets an error. Explain why.

3a. Provide the arguments for the two **set-cdr!** operations in the blanks below to produce the indicated effect on **list1** and **list2**. Do not create any new pairs; just rearrange the pointers to the existing ones.

3b. After filling in the blanks in the code above and producing the specified effect on list1 and list2, draw a box-and-pointer diagram that explains the effect of evaluating the expression (set-car! (cdr list1) (cadr list2)).

4. Exercises 3.13 and 3.14 in Abelson and Sussman.

CS 61A Week 10 Lab

Monday afternoon, Tuesday, or Wednesday morning

1. For this lab you'll need three people, each on a separate workstation.

One person should choose to be the server, and should do this:

- > (load "~cs61a/lib/im-server.scm")
- > (im-server-start)

Make a note of the IP address and port number that this prints!

The other people will be clients. They should do this:

- > (load "~cs61a/lib/im-client.scm")
- > (im-enroll "123.45.67.89" 6543) ; use actual numbers from server!

but using the server's IP address instead of 123.45.67.89 and the server's port number instead of 6543. (Note that the IP address must be enclosed in quotation marks.)

The clients can then send each other messages:

> (im 'cs61a-xy "Hi there, how are you?")

The messages can't include more than one line.

A client can leave the IM system by running

> (im-exit)

The server can quit (which disconnects all the clients) with

- > (im-server-close)
- 2. Start on the homework, since it'll be easier to test your work with several people.

CS 61A Week 11 Lab

Monday afternoon, Tuesday, or Wednesday morning

- 1. What is the type of the value of (delay (+ 1 27))? What is the type of the value of (force (delay (+ 1 27)))?
- 2. Evaluation of the expression

```
(stream-cdr (stream-cdr (cons-stream 1 '(2 3)))) produces an error. Why?
```

3. Consider the following two procedures.

What's the difference between the following two expressions?

```
(delay (enumerate-interval 1 3))
(stream-enumerate-interval 1 3)
```

4. An unsolved problem in number theory concerns the following algorithm for creating a sequence of positive integers $s_1, s_2, ...$

```
Choose s_1 to be some positive integer.
For n > 1,
if s_n is odd, then s_{n+1} is 3s_n + 1;
if s_n is even, then s_{n+1} is s_n/2.
```

No matter what starting value is chosen, the sequence always seems to end with the values 1, 4, 2, 1, 4, 2, 1, ... However, it is not known if this is always the case.

- 4a. Write a procedure num-seq that, given a positive integer n as argument, returns the stream of values produced for n by the algorithm just given. For example, (num-seq 7) should return the stream representing the sequence 7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, ...
- 4b. Write a procedure seq-length that, given a stream produced by num-seq, returns the number of values that occur in the sequence up to and including the first 1. For example, (seq-length (num-seq 7)) should return 17. You should assume that there is a 1 somewhere in the sequence.

CS 61A Week 12 Lab

Monday afternoon, Tuesday, or Wednesday morning

- 1. List all the procedures in the metacircular evaluator that call eval.
- 2. List all the procedures in the metacircular evaluator that call apply.
- 3. Explain why make-procedure does not call eval.
- 4. Abelson and Sussman, exercises 4.1, 4.2, 4.4, 4.5
- 5. In this lab exercise you will become familiar with the Logo programming language, for which you'll be writing an interpreter in project 4.

To begin, type logo at the Unix shell prompt — **not** from Scheme! You should see something like this: Welcome to Berkeley Logo version 5.5

The question mark is the Logo prompt, like the > in Scheme. (Later, in some of the examples below, you'll see a > prompt from Logo, while in the middle of defining a procedure.)

5a. Type each of the following instruction lines and note the results. (A few of them will give error messages.) If you can't make sense of a result, ask for help.

```
print 2 + 3
                                                 second "something
print 2+3
                                                 print second "piggies
print sum 2 3
                                                 pr second [another girl]
print (sum 2 3 4 5)
                                                 pr first second [carry that weight]
                                                 pr second second [i dig a pony]
print sum 2 3 4 5
                                                 to pr2nd :thing
2+3
                                                 print first bf :thing
print "yesterday
                                                 pr2nd [the 1 after 909]
print "julia"
                                                 print first pr2nd [hey jude]
print revolution
                                                 repeat 5 [print [this boy]]
print [blue jay way]
                                                 if 3 = 1+1 [print [the fool on the hill]]
show [eight days a week]
                                                 print ifelse 2=1+1 ~
show first [golden slumbers]
                                                        [second [your mother should know]] ~
                                                        [first "help]
print first bf [she loves you]
                                                 print ifelse 3=1+2 ~
pr first first bf [yellow submarine]
                                                        [strawberry fields forever] ~
                                                        [penny lane]
to second :stuff
output first bf :stuff
                                                 print ifelse 4=1+2 ~
                                                        ["flying] ~
end
                                                        [[all you need is love]]
```

```
Week 12 continued...
to greet :person
                                                 to countdown :num
say [how are you,]
                                                 if :num=0 [print "blastoff stop]
end
                                                 print :num
                                                 countdown:num-1
to say :saying
                                                 end
print sentence :saying :person
end
                                                 countdown 5
greet "ringo
                                                 to downup :word
                                                 print :word
show map "first [paperback writer]
                                                 if emptyp bl :word [stop]
                                                 downup bl :word
show map [word first ? last ?] ~
                                                 print :word
         [lucy in the sky with diamonds]
                                                 end
to who :sent
                                                 downup "rain
foreach [pete roger john keith] "describe
                                                 ;;;; The following stuff will work
                                                 ;;;; only on an X workstation:
to describe :person
print se :person :sent
                                                 cs
end
                                                 repeat 4 [forward 100 rt 90]
who [sells out]
                                                 cs
print :bass
                                                 repeat 10 [repeat 5 [fd 150 rt 144] rt 36]
make "bass "paul
                                                 cs repeat 36 [repeat 4 [fd 100 rt 90]
print :bass
                                                               setpc remainder pencolor+1 8
                                                               rt 101
print bass
                                                 to tree :size
to bass
                                                 if :size < 3 [stop]
output [johnny cymbal]
                                                 fd :size/2
end
                                                 lt 30 tree :size*3/4 rt 30
                                                 fd:size/3
print bass
                                                 rt 45 tree :size*2/3 lt 45
                                                 fd:size/6
print:bass
                                                 bk :size
                                                 end
print "bass
                                                 cs pu bk 100 pd ht tree 100
```

5b. Devise an example that demonstrates that Logo uses dynamic scope rather than lexical scope. Your example should involve the use of a variable that would have a different value if Logo used lexical scope. Test your code with Berkeley Logo.

5c. Explain the differences and similarities among the Logo operators " (double-quote), [] (square brackets), and : (colon).

CS 61A Week 13 Lab Monday afternoon, Tuesday, or Wednesday morning

For this lab, we'll be using MapReduce, a programming paradigm developed by Google that uses higherorder functions to allow a programmer to process large amount of data in parallel on many computers. Hadoop is an open source implementation of the mapreduce design.

Any computation in mapreduce consists primarily of two functions: the *mapper* and the *reducer*. (Note: The Google mapreduce paper in the course reader says "the map function" to mean the function that the user writes, the one that's applied to each datum; this usage is confusing since everyone else uses "map" to mean the higher-order function that controls the invocation of the user's function, so we're calling the latter the *mapper*:

(map mapper data)

Similarly, we'll use **reduce** to refer to the higher-order function, and **reducer** to mean the user's accumulation function.)

The mapper function takes a (input-key . input-value) pair as its argument and returns a list of (finalkey . intermediatevalue) pairs. (It returns a list of pairs, not a single pair, both to allow more than one intermediate value per input value (e.g., separating a line into words) and to allow for the possibility of returning an empty stream, meaning no intermediate values at all, so that the mapper can also effectively filter the input data.)

The reducer function is applied to each group of **intermediate-values** with the same **final-keys**, and it returns the final accumulated value. (The reducer takes two values as arguments: one new value and one partially accumulated value, just like the two-argument function used with **accumulate**.) Since there are many groups, the outputs of the reducer function are appended together to form the final output stream.

We've developed a system that allows you to run mapreduce computations on a computer cluster from within STk. Our version of mapreduce in Scheme uses slightly simpler semantics than the original MapReduce.

The syntax for performing a mapreduce computation in STk is as follows:

(mapreduce <mapper> <reducer> <base-case> <input>)

Mapper is a one-argument procedure that specifies the function that map applies.

Reducer is a procedure specifying the reduction function. Base-case is the base case for the reduction function. It is similar to the base case argument in Scheme's accumulate.

Input specifies the input data to the MapReduce computation. This argument may be a string, the name of a distributed file directory stored on all the machines in the parallel cluster, e.g., "/gutenberg" for the Project Gutenberg collection of public domain books. Alternatively, it may be a stream (not the name of a stream!) that was produced by an earlier invocation of mapreduce.

Week 13 continued:

If you forget to save the output of mapreduce, you can always run (get-last-mapreduce-output), which returns the last stream mapreduce returned.

For example, suppose we want to count the number of lines in the collected works of Shakespeare. Our input would be a set of key-value pairs with the name of a play as the key and a line of text as the value. The input is provided in a distributed file directory named "/gutenberg/shakespeare".

We could solve this problem with mapreduce as follows:

Since we're trying to get a total count for all the works of Shakespeare, not a separate count for each play, we give every intermediate pair the same key. We used the word line, but anything would have worked. The value is 1 because each line counts as one line!

What we want the reducer to do is add up all the 1s from all the files. So we don't need a complicated reducer; we just use +.

In this case, there's only one instance of **reduce** adding up all the values, but in more general cases there'll be one per key, and so what **mapreduce** returns is not a single reduced value, but a stream of key-value pairs. In this case it'll be a stream of length 1:

```
((line . number))
```

To get just the number, we'd say (kv-value (stream-car (mapreduce ...))).

Exercises:

1. The example above is inefficient because the map phase happens in parallel, but the reduce phase happens on a single machine, since all the keys are the same, and each group of same-key pairs go to a single **reduce** instance. Fix this example so that the plays are line-counted in parallel, but we still get a single total line count at the end. (Hint: Do something to the value that **mapreduce** returns.)

Week 13 continued:

2(a). One common MapReduce application is a distributed word count. Given a large body of text, such as Project Gutenberg, we want to find out which words are the most common.

Write a mapreduce program that returns each word in a body of text paired with the number of times it is used. Remember, the argument to your mapper function is a pair whose key is the document or book name and whose value is a single line of text from a single book. For example,

```
> (define gutenberg-wordcounts (mapreduce ...))
> (ss gutenberg-wordcounts 3)
((the . 300) (was . 249) (thee . 132) ... ); These aren't the actual numbers.
```

- 2(b). Using the output of the program you wrote in part (a), return the most commonly used word.
- 2(c). Using the output of the program you wrote in part (a), generate a stream of all words used only once.
- 3(a). Pattern matching is important in many areas of computer science. In this problem, we'll use mapreduce to find lines that match a pattern. Assume you're given the function match? that takes two arguments, a pattern and a sentence, and returns #t if the sentence contains a match for the pattern, or false otherwise. For example:

```
> (match? '(* hard * night *) '(a hard days night))
#t
> (match? '(* hard night *) '(a hard days night))
#f
> (match? '(* night *) '(knight rider))
#f
> (match? '(* walrus) '(I am the walrus))
#t
```

Write a procedure that takes a pattern and a directory name as arguments, and returns a stream of lines (with their keys attached) that match the pattern.

3(b). Try your pattern recognition program on the Project Gutenberg text with patterns to find some of these phrases originally attributed to Shakespeare:

pomp and circumstance one fell swoop
foregone conclusion seen better days
full circle it smells to heaven
the makings of a sorry sight
method in the madness a spotless reputation
neither rhyme nor reason strange bedfellows

CS 61A Week 14 Lab

Monday afternoon, Tuesday, or Wednesday morning

- 1. Lazy evaluator. Abelson and Sussman, exercises 4.27 and 4.29.
- 2. In this exercise we learn what a *continuation* is. Suppose we have the following definition:

```
(define (square x cont)
  (cont (* x x)))
```

Here x is the number we want to square, and **cont** is the procedure to which we want to pass the result. Now try these experiments:

```
> (square 5 (lambda (x) x))
> (square 5 (lambda (x) (+ x 2)))
> (square 5 (lambda (x) (square x (lambda (x) x))))
> (square 5 display)
> (define foo 3)
> (square 5 (lambda (x) (set! foo x)))
> foo
```

Don't just type them in – make sure you understand why they work! The nondeterministic evaluator works by evaluating every expression with two continuations, one used if the computation succeeds, and one used if it fails.

CS 61A Week 15 Lab Monday afternoon, Tuesday, or Wednesday morning

Abelson and Sussman, exercises 4.55 and 4.62:

4.55: Give simple queries that retrieve the following information from the data base:

All people supervised by Ben Bitdiddle;

The names and jobs of all people in the accounting division;

The names and addresses of all people who live in Slumerville.

4.62: Define rules to implement the last-pair operation of exercise 2.17, which returns a list containing the last element of a nonempty list. Check your rules on queries such as

```
(last-pair (3) ?x)
(last-pair (1 2 3) ?x)
(last-pair (2 ?x) (3))
```

Do your rules work correctly on queries such as (last-pair ?x (3))?

For the lab exercises and the homework problems that involve writing queries or rules, test your solutions using the query system. To run the query system and load in the sample data:

```
stk
```

```
(load "~cs61a/lib/query.scm")
(initialize-data-base microshaft-data-base)
(query-driver-loop)
```

You're now in the query system's interpreter. To add an assertion:

```
(assert! (foo bar))
```

To add a rule:

```
(assert! (rule (foo) (bar)))
```

Anything else is a query.