CSE 305 Introduction to programming languages
Spring 2016
Homework 3: Language interpreter design, part II

Assigned: March 24th, 2016 Due: on or before Monday, April 11th 11:59 pm

Overview

The goal of this homework is to understand and build an interpreter in three languages (Python, SML, Java, 20 marks for each) for a small language. Your interpreter should read in an input file (`input.txt`) which contains lines of expressions, evaluate them and push the results onto a stack, then print the content of the stack to an output file (`output.txt`) when exit. In the description below, added functionality of HW3, above that of HW2, is highlighted in yellow.

Example input and output

Some examples of the input file your interpreter takes in and the corresponding output file are shown below:

input.txt	output.txt
push 1 quit	1
push 5 Neg push 10 push 20 Add Quit	30 -5
push 10 push 2 push 8 mul add push 3 sub quit	23
push 6 push 2 div mul quit	:error: 3

input.txt	output.txt
:true: push 7 push 8 :false: pop sub quit	-1 :true:
pop push 10 swap push 6 add pop push -20 mul push 5 swap quit	-120 5 :error: 10

Functionality

Your interpreter should read in expressions from the input file named "`input.txt`", maintain a stack when running, and output the content of the stack to an output file named "`output.txt`" when stops. It should be able to handle the following expressions for this homework:

1. push

`push _num_`

where `_num_` is an integer possibly with a '-' suggesting a negative value. Here, '-0' should be regarded as '0'. Entering this expression will simply push `_num_` onto the stack. For example,

```
push 5
push -0
```

0
5

If `_num_` is not an integer, only push the error literal (`:error:`) onto the stack instead of pushing `_num_`. For example,

```
push 5
push 2.5
```

:error:
5

Pushing strings to stack:

push _string_literal_

where _string_literal_ consists of a sequence of characters enclosed in double quotation marks, as in “this is a string”. Entering this expression, would push the string onto the stack for example,

push “deadpool”
push “batman”

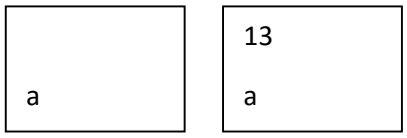
You can assume that the string value would always be legal i.e double quotes will not appear inside a string.

Pushing names to stack:

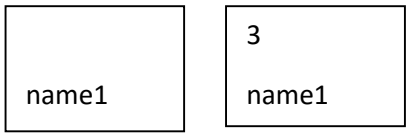
push _name_

where _name_ consists of a sequence of letters and digits, starting with a letter.

push a
push 13



push name1
push 3



To bind ‘a’ to the value 13 and name1 to the value 3, we will use ‘bind’ operation which we will see later (section 18)
You can assume that name will not contain any illegal tokens – no commas, quotation marks etc. It will always be a sequence of letters and digits starting with a letter.

2. pop

`pop`

Remove the top value from the stack. If the stack is empty, an error literal (`:error:`) will be pushed onto the stack.

For example,

3. boolean

```
:true:
:false:
```

There are two kinds of boolean literals: `:true:` and `:false:`. Your interpreter should push the corresponding value onto the stack. For example,

4. error

```
:error:
```

Similar with boolean literals, entering error literal will push `:error:` onto the stack.

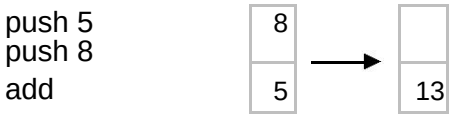
5. add

```
add
```

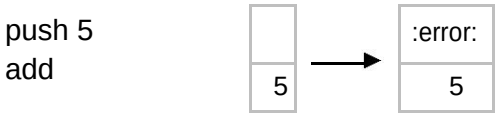
`add` refers to integer addition. Since this is a binary operator, it consumes the top two values in the stack, calculate sum and push the result back to the stack. If one of the following cases occurs, which means there is an error, any values popped out from the stack should be pushed back in the same order, then a value `:error:` should also be pushed onto the stack:

- not all top two values are integer numbers
- only one value in the stack
- stack is empty

For example,



For another example, if there is only one number in the stack and we use `add`, an error will occur. Then 5 should be pushed back as well as `:error:`



6. sub

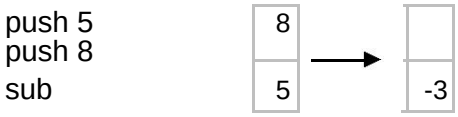
```
sub
```

`sub` refers to integer subtraction. It is a binary operator and works in the following way:

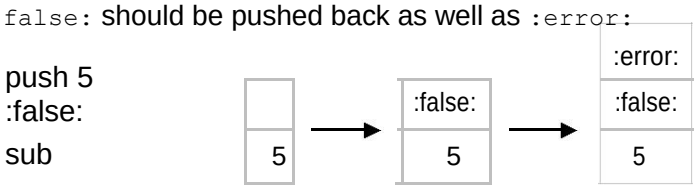
- if top two elements in the stack are integer numbers, pop the top element(y) and the next element(x), subtract y from x, and push the result x-y back onto the stack
- if the top two elements in the stack are not all integer numbers, push them back in the same order and push `:error:` onto the stack

- if there is only one element in the stack, push it back and push `:error:` onto the stack
- if the stack is empty, push `:error:` onto the stack

For example,



For another example, if one of the top two values in the stack is not a numeric number when `sub` is used, an error will occur. Then 5 and `:false:` should be pushed back as well as `:error:`



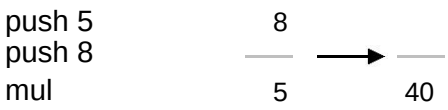
7. mul

`mul`

`mul` refers to integer multiplication. It is a binary operator and works in the following way:

- if top two elements in the stack are integer numbers, pop the top element(y) and the next element(x), multiply x by y, and push the result x*y back onto the stack
- if the top two elements in the stack are not all integer numbers, push them back in the same order and push `:error:` onto the stack
- if there is only one element in the stack, push it back and push `:error:` onto the stack
- if the stack is empty, push `:error:` onto the stack

For example,



For another example, if the stack empty when mul is used, an error will occur. Then :error: should be pushed onto the stack:

8. div

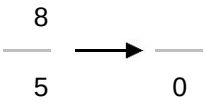
`div`

`div` refers to integer division. It is a binary operator and works in the following way:

- if top two elements in the stack are integer numbers, pop the top element(y) and the next element(x), divide x by y, and push the result x\y back onto the stack
- if top two elements in the stack are integer numbers but y equals to 0, push them back in the same order and push `:error:` onto the stack
- if the top two elements in the stack are not all integer numbers, push them back in the same order and push `:error:` onto the stack
- if there is only one element in the stack, push it back and push `:error:` onto the stack
- if the stack is empty, push `:error:` onto the stack

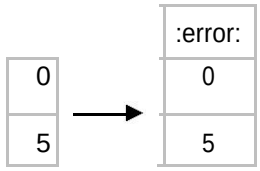
For example,

```
push 5
push 8
div
```



For another example, if the top element in the stack equals to 0, there will be an error if `div` is used. Then 5 and 0 should be pushed back as well as `:error:`

```
push 5
push 0
div
```



9. rem

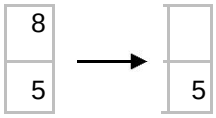
`rem`

`rem` refers to the remainder of integer division. It is a binary operator and works in the following way:

- if top two elements in the stack are integer numbers, pop the top element(y) and the next element(x), calculate the remainder of x\y, and push the result back onto the stack
- if top two elements in the stack are integer numbers but y equals to 0, push them back in the same order and push `:error:` onto the stack
- if the top two elements in the stack are not all integer numbers, push them back and push `:error:` onto the stack
- if there is only one element in the stack, push it back and push `:error:` onto the stack
- if the stack is empty, push `:error:` onto the stack

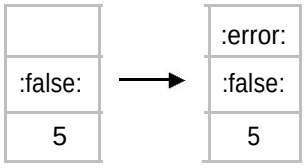
For example,

```
push 5
push 8
rem
```



For another example, if one of the top two elements in the stack is not an integer, an error will occur if `rem` is used. Then 5 and `:false:` should be pushed back as well as `:error:`

```
push 5
:false:
rem
```



10. neg

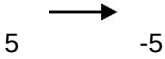
`neg`

`neg` is to calculate the negation of an integer (negation of 0 should still be 0). It is unary therefore consumes only the top element from the stack, calculate its negation and push the result back. A value `:error:` will be pushed onto the stack if:

- the top element is not an integer, push the top element back and push `:error:`
- the stack is empty, push `:error:` onto the stack

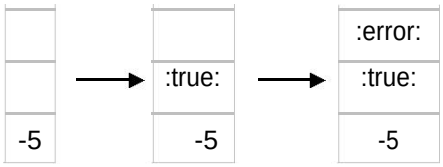
For example,

```
push 5
neg
```



For another example, if the top value is not an integer, when `neg` is used, it should be pushed back as well as `:error:`

```
push 5
neg
:true:
neg
```



11. swap

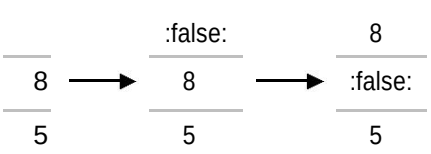
`swap`

`swap` interchanges the top two elements in the stack, meaning that the first element becomes the second and the second becomes the first. A value `:error:` will be pushed onto the stack if:

- there is only one element in the stack, push the element back and push `:error:`
- the stack is empty, push `:error:` onto the stack

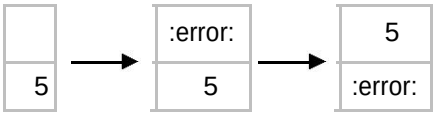
For example,

```
push 5
push 8
:false:
swap
```



For another example, if there is only one element in the stack when `swap` is used, an error will occur and `:error:` should be pushed onto the stack. Now we have two elements in the stack (5 and `:error:`), therefore the second swap will interchange the two elements:

```
push 5
swap
swap
```



12. quit

`quit`

`quit` causes the interpreter to stop. Then the whole stack should be printed out to an output file, named as "`output.txt`".

13. and

`and`

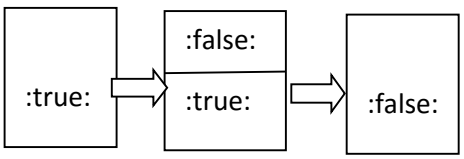
`and` performs the logical conjunction of the top two elements in the stack and pushes the result (a single value) onto the stack.

`:error:` will be pushed onto the stack if:

- there is only one element in the stack, push the element back and push `:error:`
- the stack is empty, push `:error:` onto the stack
- if either of the top two elements aren't Boolean, push back the elements and push `:error:`

For example,

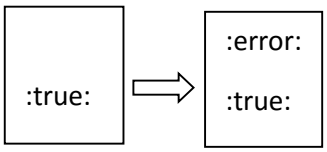
```
:true:
:false:
and
```



Another example,

:true:

and



14. or

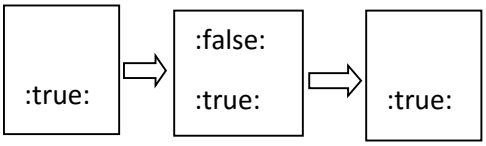
`or`

`or` performs the logical disjunction of the top two elements in the stack and pushes the result (a single value)onto the stack.

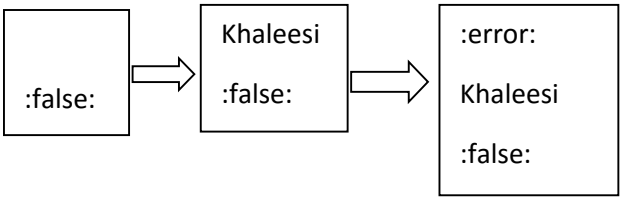
`:error:` will be pushed onto the stack if:
- there is only one element in the stack, push the element back and push `:error:`
- the stack is empty, push `:error:` onto the stack
- if either of the top two elements aren't Boolean, push back the elements and push `:error:`

For example,

```
:true:
:false:
or
```




```
:false:
push "Khaleesi"
or
```



15. not

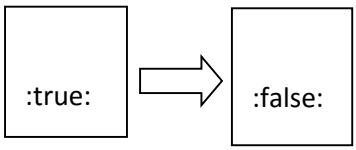
`not`

`not` performs the logical negation of the top element in the stack and pushes the result (a single value)onto the stack. Since the operator is unary, it only consumes the top value from the stack.

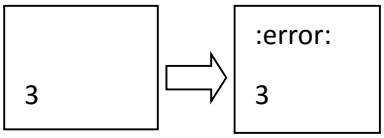
`:error:` will be pushed onto the stack if:
- the stack is empty, push `:error:` onto the stack
- if the top element isn't Boolean, push back the element and push `:error:`

For example,

:true:
not



push 3
not



16. equal

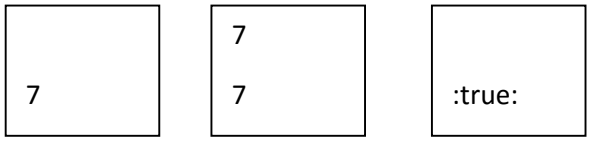
`equal`

`equal` refers to numeric equality (so you are not supporting string comparisons) This operator consumes the top two values on the stack and pushes the result(a single boolean value) onto the stack.

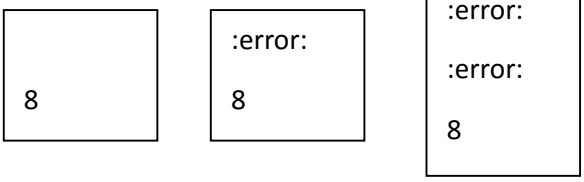
`:error:` will be pushed onto the stack if:
- there is only one element in the stack, push the element back and push `:error:`
- the stack is empty, push `:error:` onto the stack
- if either of the top two elements aren't integers, push back the elements and push `:error:`

For example,

push 7
push 7
equal



push 8
push 9.5
equal



17. lessThan

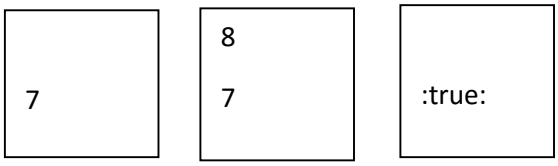
`lessThan`

`lessThan` refers to numeric less than ordering. This operator consumes the top two values on the stack and pushes the result(a single Boolean value) onto the stack.

`:error:` will be pushed onto the stack if:
- there is only one element in the stack, push the element back and push `:error:`
- the stack is empty, push `:error:` onto the stack
- if either of the top two elements aren't integers, push back the elements and push `:error:`

For example,

push 7
push 8
lessThan



18.bind

`bind`

`bind` binds a name to a value. It is evaluated by popping two values from the stack. The second value popped must be a name (see section on push for details on what constitutes a 'name'). The name is bound to the value (the first thing popped off the stack). The value can be any of the following :

- An integer
- A string
- Boolean
- :unit:
- The value of a name that has been previously bound

The name value binding is stored in an environment data structure. The result of a bind operation is :unit: which is pushed onto the stack.

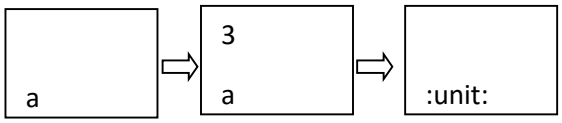
`:error:` will be pushed onto the stack if:

Modification 3/26/16

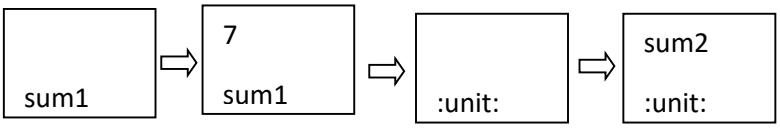
- ~~the name is already bound in the current environment~~
- If we are trying to bind an identifier to an unbound identifier.
- the stack is empty, push `:error:` onto the stack
  in which case all elements popped must be pushed back before pushing :error: onto the stack.

For example :

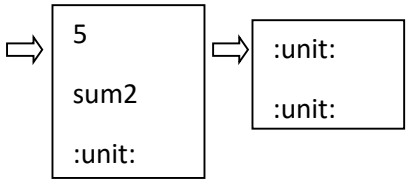
push a
push 3
bind



push sum1
push 7
bind

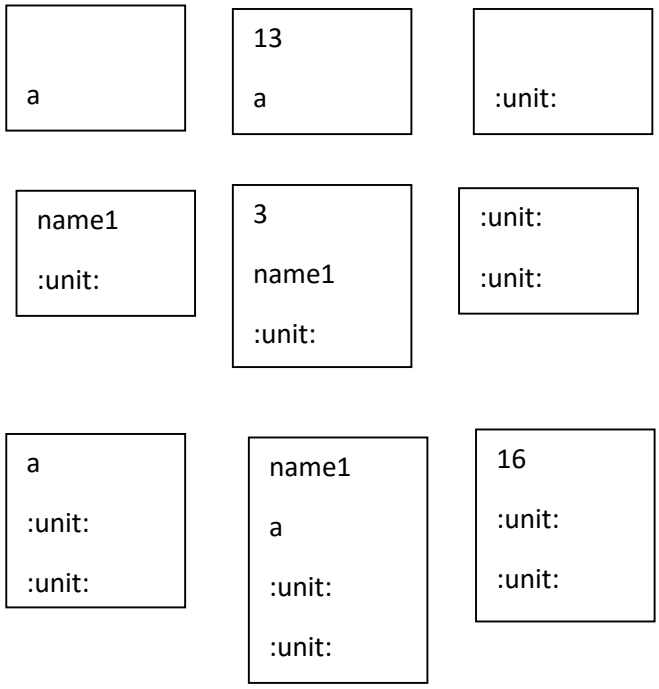


```
push sum2
push 5
bind
```



You can use bindings to hold values which could be later retrieved and used by functionalities you already implemented. For instance in the example below, an addition on a + name1 in example1, would add 13 + 3 and push the result 16 onto the stack.

```
push a
push 13
bind
push name1
push 3
bind
push a
push name1
add
```



While performing operations, if a name has no binding, push :error: onto the stack.

Bindings can be overwritten, for instance:

```
push a
push 9
bind
push a
push 10
bind
```

Here, the second bind updates the value of ‘a’ to 10.

19.if

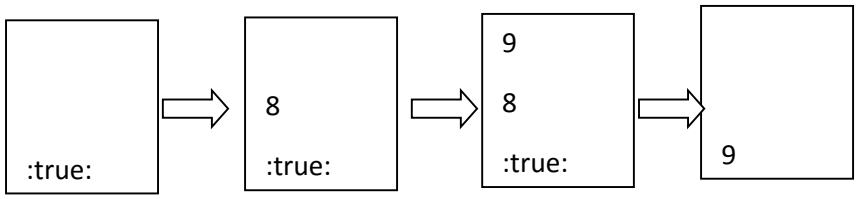
`if`

`if` pops three values off the stack; x,y and z. The third value popped (z, in this case) must always be a Boolean. If z is :true:, if has the value of x, and if z is :false:, if has the value y.

`:error:` will be pushed onto the stack if:

- the third value is not Boolean.
- the stack is empty, push `:error:` onto the stack
- there are less than 3 values on the stack in which case all elements popped must be pushed back before pushing :error: onto the stack.

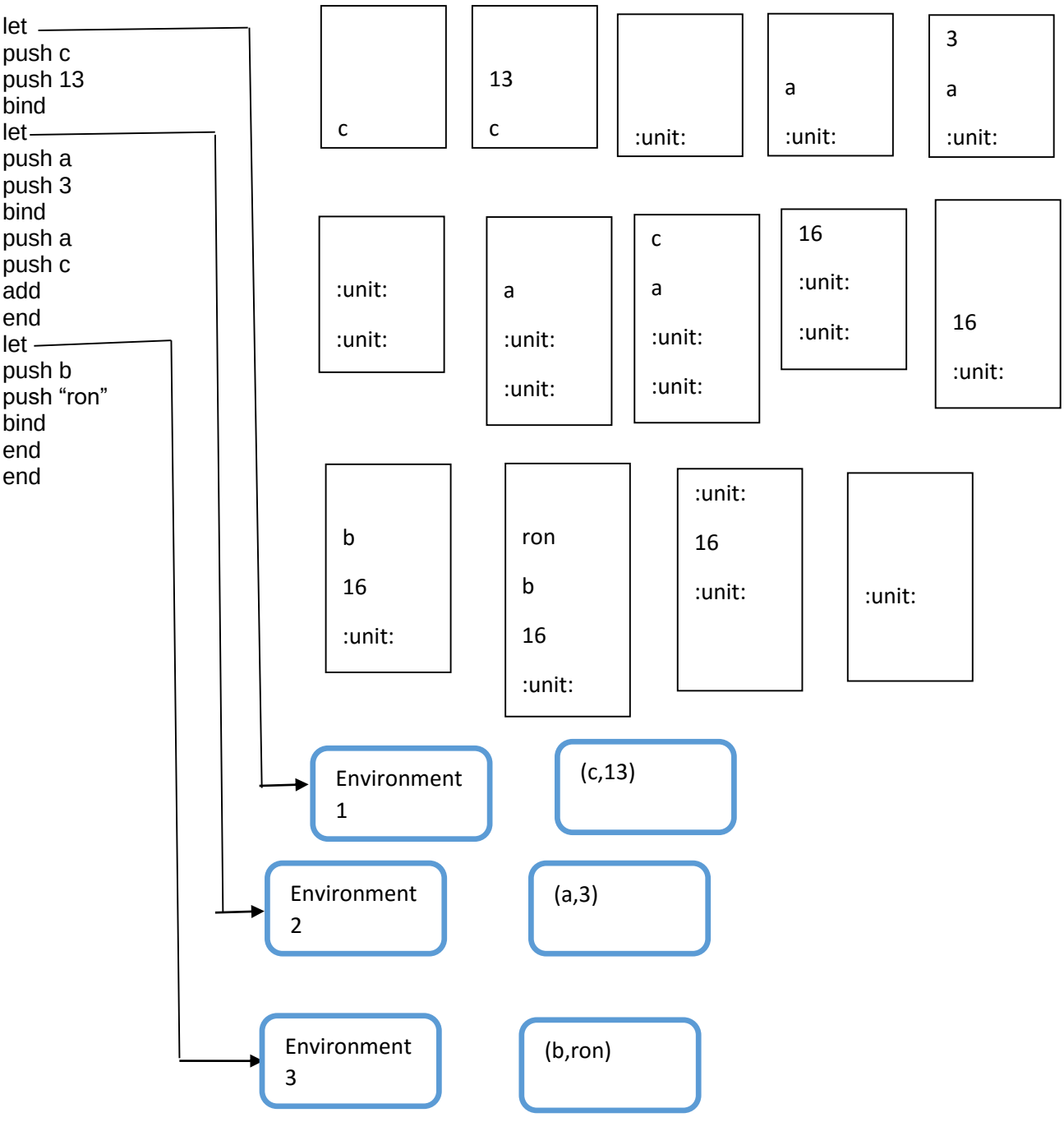
:true:
push 8
push 9
if



20. let...end

`let...end` limits the scope of variables. "let" marks the beginning of a new environment – which is basically a sequence of bindings. The result of the let..end is the last stack frame of the let. Let..end can contain any number of operations but it will always result in a stack frame that is strictly larger than the stack prior to the let.

Trying to access an element that is not in scope of the let..end block would push :error: on the stack. let..end blocks can also be nested.



In the above example, the first let statement creates an empty environment (environment 1), then the name c is bound to 13. The result of this bind is a :unit: on the stack and a name value pair in the environment. The second let statement creates a second empty environment. Name a is bound here. To add a and c, these names are first looked up for their values in the current environment. If the value isn't found in the current environment, it is searched in the outer environment. Here, c is found from environment 1. The sum is pushed to the stack. A third environment is created with one binding 'b'.The second last end is to end the scope of environment 3 and the last end statement is to end the scope of environment 1.

You can assume that the stack is left with at least 1 item after the execution of any let..end block.

You can make the following assumptions:

- Expressions given in the input file are in correct formats. For example, there will not be expressions like "`push`", "`3`" or "`add 5`"
- No multiple operators in the same line in the input file. For example, there will not be "`pop pop swap`", instead it will be given as

```
pop
pop
swap
```

- There will always be a "`quit`" in the input file to exit your interpreter and output the stack

You can still assume that all test cases will have a quit statement at the end.

You can assume that your hw3 function will only be called ONCE per execution of your program.

Step by step examples

If your interpreter reads in expressions from “input.txt”, states of the stack after each operation are shown below:

input.txt
push 10
push 15
push 30
sub
:true:
swap
add
pop
neg
quit

First, push 10 onto the stack:

10

Similarly, push 15 and 30 onto the stack:

30
15
10

`sub` will pop the top two values from the stack, calculate 15-30 = -15, and push -15 back:

-15
10

Then push the boolean literal `:true:` onto the stack:

:true:
-15
10

`swap` consumes the top two values, interchanges them and pushes them back:

-15
:true:
10

`add` will pop the top two values out, which are -15 and `:true:`, then calculate their sum. Here, `:true:` is not a numeric value therefore push both of them back in the same order as well as an error literal `:error:`

:error:
-15
:true:
10

`pop` is to remove the top value from the stack, resulting in:

-15
:true:
10

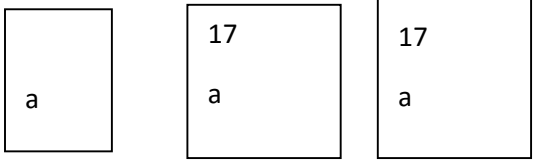
Then after calculating the negation of -15, which is 15, and pushing it back, `quit` will terminate the interpreter and write the following values in the stack to "`output.txt`":

15
:true:
10

Now, please go back to the example inputs and outputs given before and make sure you understand how to get those results.

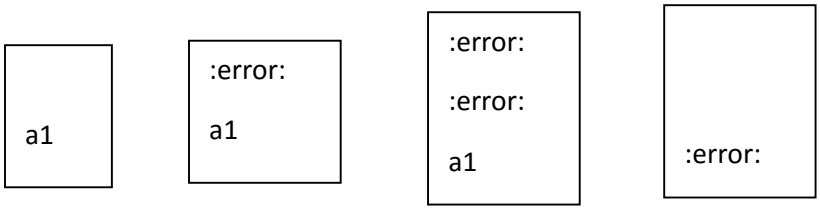
More Examples of bind and let..end:

```
push a
push 17
add
```



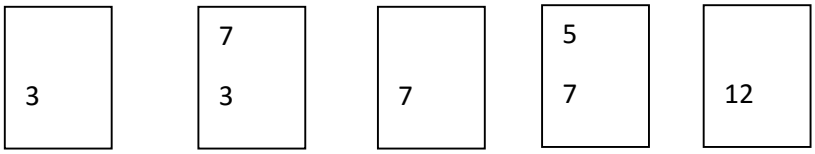
The error is because we are trying to perform an addition on an unbound variable "a".

```
let
push a1
push 7.2
bind
end
```



Additional examples for let..end and frequently asked questions:

```
let
push 3
push 7
end
push 5
add
quit
```



Explanation :

Push 3
Push 7

Pushes 3 and 7 on top of the stack. When you encounter the "end", the last stack frame is saved (which is why the value of 7 is retained on the stack) , then 5 is pushed onto the stack and the values are added.

You may ask - But isn't the 7 local to the let..end? 7 is not a binding – it is just a value. The local scopes are only for bindings.

FAQs from Piazza:

1. What values can _name_ be bound to?

   name can be bound to integers, Boolean, string, :unit: and also previously bound values. For example,

   ```
   push a
   :true:
   bind
   ```

   would bind a to :true:

   ```
   push a
   push 7.5
   bind
   ```

   would result in bind producing an :error: because a CANNOT be bound to :error:

   ```
   push b
   ```

```
let
push a
push 7
bind
end
bind
```

would bind a to 7 and b to :unit:

```
push b
push 8
bind
push a
push b
bind
```

would bind b to 8 and would bind a to the VALUE OF b which is 8.

```
push b
push a
bind
```

would result in an :error: because you are trying to bind b to an unbound variable a.

2. What values can ‘if’ take?

The result of executing a ‘if’ can be an integer or Boolean or string or :error: or :unit:

For instance,

```
:true:
push “oracle”
push “jive”
if
```

the result of if would be “jive”

:false:

```
let
push a
push 8
bind
end
push 8.9
if
```

the result of if would be :unit:

3. What is the result of executing the following:

```
push a
push 5
bind
pop
:true:
push 4
push a
if
```

The stack would have a. Although the value of a is bound to 5, we only resolve the name to the value if we need to perform computation. (For 'if', the only value needed for computation is Boolean.)

4. What would be the result of executing the following :

```
let
push a1
push 7.2
bind
end
quit
```

7.2 cant be pushed to the stack and a1 cannot be bound to :error: so, the result would be :error:

5. How can we bind identifiers to previously bound values?

```
push a
push 7
bind
push b
push a
bind
```

The first bind binds the value of a to 7. The second bind statement would result in the name b getting bound to the VALUE of a – which is 7. This is how we can bind identifiers to previously bound values. Note that we are not binding b to a – we are binding it to the VALUE of a.

6. What would be the output of running the following :

```
push 1
let
push 2
push 3
push 4
end
push 5
```

This would result in the stack :

```
5
4
1
```

Explanation : After the let..end is executed the last frame is returned – which is hy we have 4 on the stack.

7. Can we have something like this:

```
push a
push 15
push a
```

Yes. In this case ‘a’ is not bound to any value yet. And the stack contains:

```
a
15
a
```

If we had :

```
push a
push 15
bind
push a
```

The stack would be :

```
a
:unit:
```

8. What would be the output of running the following:

```
let
push 3
end
let
push b
```

swap
bind
end

The stack would result in :unit:
(3 is a value – not a binding and hence is not limited to the scope of the first let..end)
We will NOT be testing code like this since this violates the assumption that let..end is monotonically increasing. So we do NOT expect your code to handle such cases.

9. What would be the output of running the following code:

```
let
push 3
push 10
end
add
quit
```

The stack output would be
:error:
10

10. Can we push the same _name_ twice to the stack? For instance , what would be the result of the following:

```
push a
push a
quit
```

This would result in the following stack output:
a
a

Yes, you can push the same _name_ twice to the stack. Consider binding it this way :

```
push a
push a
push 2
bind
```

This would result in
:unit: → as a result of binding a to 2
a → as a result of pushing the first a to the stack

11. Output of the following code:

```
push a
push 9
bind
push a
```

push 10
bind

would result in
:unit: → as a result of second bind
:unit: → as a result of first bind

IMPORTANT PLEASE READ:

What to submit

Define a function named 'hw3', which takes in two strings as input arguments (first one is the name of an input file and the second is the name of an output file). Function signatures of 'hw3' will be the same as they are specified in homework 1 **(please see instructions in 'HW1_Modified.pdf').**

Create a folder `UBITName_HW3` which contains three sub folders: `Python`, `SML` and `Java`. The '`Python`' folder should contain **ONLY** '`hw3.py`', '`SML`' folder should contain **ONLY** '`hw3.sml`', and '`Java`' folder should contain **ONLY** '`hw3.java`'. **Please DO NOT submit any input output files or test scripts.**

Compress the `UBITName_HW3` folder to either '`UBITName_HW3.zip`' or '`UBITName_HW3.tar`', and submit it on Timberlake using the command `submit_cse305 your_file_name`.

Auto-grader will be used to grade your submission, so please follow the exact naming conventions mentioned above. If your code fails the auto grader due to incorrect naming conventions, you will receive a 0.

No late submissions will be accepted (unless you decide to use your "free days"). So please start early.

BEFORE YOU SUBMIT :

- **Did you name your function hw3() – for the SML, Java and Python part?**
- **Did you remove the function call to hw3() from the Python and SML part?**

- **Did you remove the Tester file/ class file/ input output files/ any extra files from the Python, Java and SML sub folders?**
- **Is your home folder called yourUBITName_HW3 ?**
- **Does it have ONLY three sub folders in it called Java, Python and SML?**
- **Does the Java folder have JUST ONE FILE called hw3.java?**
- **Does the SML folder have JUST ONE FILE called hw3.sml?**
- **Does the Python folder have JUST ONE FILE called hw3.py?**
- **Did you zip/ tar your home folder? (NO tar.gz or any other forms of compression please)**
- **Before you submit, did you unzip/untar it and double check to make sure you aren't making an empty file submission?**