CS1101S Discussion Group Week 5:
Data Abstraction & List Processing

Niu Yunpeng

niuyunpeng@u.nus.edu

September 12, 2017

Overview

1. Data abstraction
   - What is data
   - To understand data structure
   - To use data structure
2. Pair & list
   - Pair processing
   - An "insider" problem
   - List processing
   - Exercises

Data Abstraction

What is data?

- Data is the storage of information.
- Two kinds of information: **states** & **procedures**.
- Procedures are the manipulation of states.

Data in the Source

- *To represent states*: use variables;
- *To represent procedures*: use functions.

Data Abstraction

Variables & functions

- Variables are data;
- Functions are procedures.
- Meanwhile, procedures are also data.

Higher-order programming

- Variables can be functions.
- Parameters can be functions.
- Return values can be functions.

Data Abstraction

Still remember highest_denom() in lecture notes?

```
function highest_denom(kind) {
    if (kind === 1) {
        return 5;
    } else if (kind === 2) {
        return 10;
    } else if (kind === 3) {
        return 20;
    } else if (kind === 4) {
        return 50;
    } else if (kind === 5) {
        return 100;
    } else {
        display("invalid coin");
    }
}
```

Data Abstraction

What is `highest_denom()` about?

- We want to know the value for each kind of coins. We certainly can store them in variables like `coinA`, `coinB`, `coinC`, etc.
- However, what if we have too many kinds of coins? We then need a *well-organized structure* to store all the information.

What if we have too many kinds of coins?

- We then need a ***well-organized structure*** to store all the data.

Data Abstraction

What is *data structure*?

- Data structure provides us with a *well-organized* way to store all related information as a collection.
- Data structure should provide functions so that we can arbitrarily get/change the values inside.
  - `getters`
  - `setters`

Data Abstraction

Data structure & black-box abstraction

- Data structure is a black-box.
- We can use it to store and retrieve data without knowing things inside.

Data Abstraction

Use data structure with `highest_denom`

The data structure should at least provide the functions below to use:

- `initialize()`: to initialize a new data structure to store different kinds of coins and their respective values;
- `add_new_kind(id, value)`: to add a new kind of coins to an existing data structure with a unique identifier and its value;
- `get_value(id)`: to get the corresponding value of a certain kind of coins by its unique identifier.

Data Abstraction

To use data structure

- Revisit the example on the lecture notes - rationals.
- Try to understand how to *design and build* a *tailor-made data structure* for a specific problem.

Data Abstraction

Rational numbers

The data structure should at least provide the functions below to use:

- `make_rat(num, denom)`: make a rational number with its numerator and its denominator;
- `get_num(rat)`: get the numerator of a rational;
- `get_denom(rat)`: get the denominator of a rational;
- `add_rat(a, b)`: add two rationals *a* and *b*;
- `sub_rat(a, b)`: subtract two rationals *a* and *b*;
- `mul_rat(a, b)`: multiply two rationals *a* and *b*;
- `div_rat(a, b)`: make a division of two rationals *a* and *b*;
- `equal_rat(a, b)`: check whether two rationals are equal;
- `rat_to_string(rat)`: convert a rational to a string.

Data Abstraction

Make a rational number

```
function make_rat(num, denom) {
    var divider = gcd(num, denom);
    return pair(num / divider, denom / divider);
}

function get_num(rat) {
    return head(rat);
}

function get_denom(rat) {
    return tail(rat);
}
```

Data Abstraction

Rational number calculation

```
function add_rat(a, b) {
    return make_rat(get_num(a) * get_denom(b) +
                    get_num(b) * get_denom(a),
                    get_denom(a) * get_denom(b));
}

function sub_rat(a, b) {
    return make_rat(get_num(a) * get_denom(b) -
                    get_num(b) * get_denom(a),
                    get_denom(a) * get_denom(b));
}
```

Data Abstraction

Rational number calculation

```
function mul_rat(a, b) {
    return make_rat(get_num(a) * get_num(b),
                    get_denom(a) * get_denom(b));
}

function div_rat(a, b) {
    return make_rat(get_num(a) * get_denom(b),
                    get_denom(a) * get_num(b));
}
```

Data Abstraction

Others

```
function equal_rat(a, b) {
    return get_num(a) === get_num(b) &&
           get_denom(a) === get_denom(b);
}

function rat_to_string(rat) {
    return get_num(rat) + "/" + get_denom(rat);
}
```

Overview

1. Data abstraction
   - What is data
   - To understand data structure
   - To use data structure

2. Pair & list
   - Pair processing
   - An “insider” problem
   - List processing
   - Exercises

Pair & List Processing

Use `pair` as a data structure

The data structure should at least provide the functions below to use:

- `pair(x, y)`: construct a pair with two elements *a* and *b*;
- `head(some_pair)`: get the first element of a pair;
- `tail(some_pair)`: get the second element of a pair;
- `is_pair(some_pair)`: check whether an object is a pair.

Pair & List Processing

Three ways to represent a `pair`

- Use your code in the Source language;
- Use box-and-pointer diagram (as the list visualizer);
- Use square brackets (as the output in the interpreter).

Notice

- The same applies to `list` later.

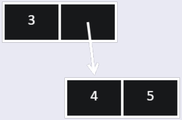
Pair & List Processing

Three ways to represent a pair

- Use your code in the Source language;
- Use box-and-pointer diagram (as the list visualizer);
- Use square brackets (as the output in the interpreter).

Example

- `var x = pair(3, pair(4, 5));`



- [3, [4, 5]]

Pair & List Processing

Consider: `make_one_out_of_two`

```
function make_one_out_of_two(a, b) {
    return function(oper) {
        return oper(a, b);
    };
}

function first(pair) {
    return pair(function(m, n) { return m; });
}

function second(pair) {
    return pair(function(m, n) { return n; });
}
```

Pair & List Processing

From pair to list

- Sometimes, we need to store more than 2 variables in a data structure.
- Without list, we have to

```
pair(3, pair(1, pair(4, pair(1, pair(5, ...)))));
```

- With list, we only need to

```
list(3, 1, 4, 1, 5, ...);
```

Pair & List Processing

Formal definition

- A list is either an empty list or a pair whose tail is a list.

Use `list` as a data structure

Up to now, we have the following functions to use:

- `list(x, y, z, ...)`: construct a list with *n* elements;
- `head(lst)`: get the first element of a list;
- `tail(lst)`: get the remaining part of a list;
- `is_list(lst)`: check whether an object is a list;
- `is_empty_list(lst)`: check whether an object is a list and empty;
- `length(lst)`: count the number of elements in a list.

Pair & List Processing

Recap: three ways to represent `pair` and `list`

- Use your code in the Source language;
- Use box-and-pointer diagram (as the list visualizer);
- Use square brackets (as the output in the interpreter).

Pair & List Processing

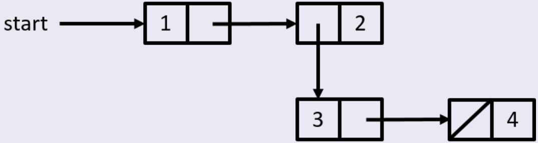
Exercise 1

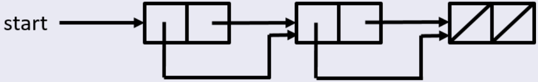
Draw the box-and-pointer diagrams for each one of them below:

```
var lstA = list(list([], 1, list([], 2, [])),
                3,
                list([], 4, []));

var p1 = pair(4, []);
var p2 = pair(3, p1);
var lstB = list(1, pair(2, p2));

var z1 = pair(1, 3);
var z2 = list(3, z1);
var lstC = list(tail(z2), z1, head(z1));
```

Exercise 2

Write Source programs which can produce the box-and-pointer diagrams below (*The head of the whole list should be pointing to "start"*):





Exercise 3

Given two lists of the same length `xs` and `ys`, try to construct a 3^{rd} list of the same length in which each element is a pair composed of the element on the same position from `xs` and `ys`. Your function name should be `make_pairs`.

Example

For example, for `make_pairs(list(1, 2, 3), list(11, 12, 13))`, it should return `list(pair(1, 11), pair(2, 12), pair(3, 13))`.

Pair & List Processing

Exercise 3

Now, generalize this concept by defining a new function. Given two lists of the same length `xs` and `ys`, try to construct a 3^{rd} list of the same length in which each element is the result of applying a certain zip function to the two elements on the same position from `xs` and `ys`. Your function name should be `zip`.

Example

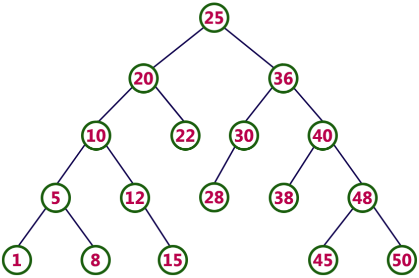
For example, if we apply

```
zip(function (x, y) { return x * y; },
    list(1, 2, 3),
    list(11, 12, 13));
```

it will return `list(11, 24, 39)`.

Exercise 4 - BST

A binary search tree (BST) is either an empty list or a list with three elements: a left child BST, a number x, and a right child BST. Notice that every number in the left BST is smaller than the number x, and every number in the right BST is larger than the number x.



Exercise 4 - BST

The first step to understand how to use BST is to have a try. Given 5 numbers 1...5, try to store them in a BST. Then, you should use the 3 ways to represent this list (notice: BST is just a special kind of list). The answer may not be unique.

Exercise 4 - BST

The data structure should at least provide the functions below to use:

- `get_min(tree)`: get the smallest element in a BST;
- `get_max(tree)`: get the largest element in a BST;
- `search(tree, x)`: check whether a number exists in a BST;
- `height(tree)`: get the height of a BST;
- `bst_to_list(tree)`: convert a BST into a list.

Task

Implement all these functions mentioned above and other necessary functions that should be supported by a BST library.

Discussion Group Problems

Let's discuss them now.

The End