ABSTRACT DATA TYPES AND IMMUTABLE TREES Solutions

COMPUTER SCIENCE MENTORS 61A

October 3–October 007, 2022

1 Abstraction

Data abstraction allows us to create and access data through a controlled, restricted programming interface—hiding implementation details for sake of brevity and reusability of code and encouraging programmers to focus on how data is used rather than worrying about how data is internally organized. The two fundamental components of an **abstract data type** are a constructor and selectors:

1. A **constructor** creates a piece of data, and includes all the attributes that make the data unique; e.g. executing `c = car("Nissan", "Leaf")` creates a new instance of a car abstraction and assigns it to the variable `c`.
2. **Selectors** access attributes of a piece of data; e.g. calling `get_make(c)` returns `"Nissan"`.

In the example above, you don't know specifically how the model name "Nissan" and the make name "Leaf" are internally bundled into a car, and you don't care, either. The creator of the abstract data type dealt with those details, so that you, the user of the ADT, would only have to know how to store and retrieve the data you need. This separation of concerns between designing and using an interface is called the **abstraction barrier**. While your program won't necessarily break if you break the abstraction barrier, heeding the barrier is best practice and can prevent errors down the road.

Using abstraction to hide unnecessary details can be seen everywhere, not just in code—keyboards, printers, cars, stovetops, and typewriters all employ abstractive interfaces. What are some examples of abstraction in your everyday life?

Vibha Tantry and Jordan Schwartz, with
Gabe Classon, Aditya Balasubramanian, Oliver Yu, Hanze Tan, Gene Pan, Carolyn Wang, Raman Varma, Noor Haq, Angela Zhang, Abigail Brooks, Ted Kim, Aiden Joshua Vehemente, Marie Chorpita, Ashley Chiu, Bill Hu, Omar Yu, Angel Alberto Aldaco, Andrew Park, Chase Clements, Grace Yi, Preshtha Garg, Laksith Venkatesh Prabu, Aditya Murali, Irene Yang, Andy Chen, and Hailey Park

1. The following is an abstract data type that represents Pokemon. Each Pokemon keeps track of its name, type, and friends. Given our provided constructor, fill out the selectors:

```
def pokemon(name, p_type, friends):
    """
    Constructs a Pokemon with the given attributes.
    >>> cyndaquil = pokemon('Cyndaquil', 'Fire',
       ['Chikorita', 'Totodile'])
    >>> p_name(cyndaquil)
    'Cyndaquil'
    >>> p_type(cyndaquil)
    'Fire'
    >>> p_friends(cyndaquil)
    ['Chikorita', 'Totodile']
    """
    return [name, p_type, friends]

def p_name(p):

    return p[0]

def p_type(p):

    return p[1]

def p_friends(p):

    return p[2]
```

2. This function returns the correct result, but there's something wrong with its implementation. What's the issue, and how can we fix it?

```python
def are_friends(p1, p2):
    """
    Returns True iff the Pokemon p1 and p2 are each other's
       friends.
    """
    return p1[0] in p2[2] and p2[0] in p1[2]
```

Treating the `p1` and `p2` are lists is a Data Abstraction Violation (DAV). We should use a selector instead. The corrected function looks like:

```python
def are_friends(p1, p2):
    return p_name(p1) in p_friends(p2) and p_name(p2) in
       p_friends(p1)
```

3. Write the function `cross_type_friends`, which takes in a Pokemon `p` and a list of Pokemon `pokemon_list` and returns a list of the names of `p`'s cross-type friends in `pokemon_list`. (A cross-type friend is a friend of a different type.) You may assume that the `are_friends` function has been correctly implemented.

```
def cross_type_friends(p, pokemon_list):
    """
    >>> c = pokemon('Charmander', 'Fire', ['Torchic',
       'Squirtle', 'Bulbasaur'])
    >>> t = pokemon('Torchic', 'Fire', ['Charmander',
       'Squirtle'])
    >>> s = pokemon('Squirtle', 'Water', ['Torchic',
       'Bulbasaur'])
    >>> b = pokemon('Bulbasaur', 'Grass', ['Charmander',
       'Squirtle'])
    >>> cross_type_friends(c, [t, s, b])
    ['Bulbasaur']
    >>> cross_type_friends(b, [c, s, b])
    ['Charmander', 'Squirtle']
    """

    friend_list = []
    for other in pokemon_list:
        if are_friends(p, other) and p_type(p) !=
           p_type(other):
            friend_list += [p_name(other)]
    return friend_list

    # Alternative solution

    return [p_name(o) for o in pokemon_list if are_friends(p,
       o) and p_type(p) != p_type(o)]
```

4. In this problem, you'll change the implementation of the Pokemon ADT while keeping the interface the same.

   (a) Complete the constructor for the given selectors.

```
def pokemon(name, p_type, friends):
    """
    >>> lil_guy = pokemon('Pikachu', 'Electric',
       ['Mewtwo', 'Lucario'])
    >>> p_name(lil_guy)
    'Pikachu'
    >>> p_type(lil_guy)
    'Electric'
    >>> p_friends(lil_guy)
    ['Mewtwo', 'Lucario']
    """

    def select(command):
        if command == 'name':
            return name
        elif command == 'type':
            return p_type
        elif command == 'friends':
            return friends
    return select
```

Alternate solution:

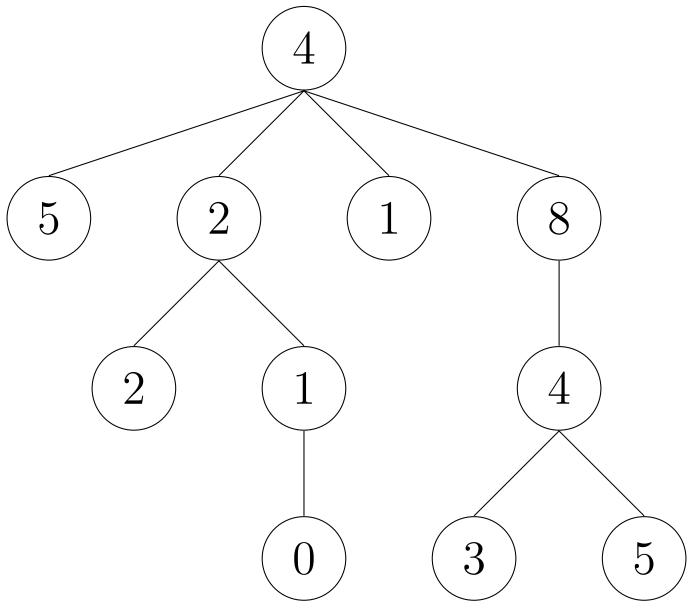
```
    return lambda sel: {'name': name, 'type': p_type,
       'friends': friends}[sel]
```

```
def p_name(p):
    return p('name')

def p_type(p):
    return p('type')

def p_friends(p):
    return p('friends')
```

(b) What do we need to change about the implementations of `are_friends` (as revised) and `cross_type_friends` now that we've changed the implementation of the Pokemon ADT? Why?

Nothing. Because we relied on the implementation-independent interface of the Pokemon ADT, changing the underlying implementation does not affect the correctness of these functions.

Trees are a kind of recursive data structure. Each tree has a **root label** (which is some value) and a sequence of **branches**. Trees are "recursive" because the branches of a tree are trees themselves! A typical tree might look something like this:



This tree's root label is 4, and it has 4 branches, each of which is a smaller tree. The 6 of the tree's **subtrees** are also **leaves**, which are trees that have no branches.

Trees may also be viewed **relationally**, as a network of nodes with parent-child relationships. Under this scheme, each circle in the tree diagram above is a node. Every non-root node has one parent above it and every non-leaf node has at least one child below it.

Trees are represented by an abstract data type with a `tree` constructor and `label` and `branches` selectors. The `tree` constructor takes in a label and a list of branches and returns a tree. Here's how one would construct the tree shown above with `tree`:

```
tree(4,
     [tree(5),
      tree(2,
          [tree(2),
           tree(1,
               [tree(0)])]),
      tree(1),
      tree(8,
          [tree(4,
               [tree(3), tree(5)])])])
```

The implementation of the ADT is provided here, but you shouldn't have to worry about this too much. (Remember the abstraction barrier!)

```
def tree(label, branches=[]):
        return [label] + list(branches)

def label(tree):
        return tree[0]

def branches(tree):
        return tree[1:] # returns a list of branches
```

Because trees are recursive data structures, recursion tends to a be a very natural way of solving problems that involve trees.

- The **recursive case** for tree problems often involves recursive calls on the branches of a tree.
- The **base case** is often reached when we hit a leaf because there are no more branches to recurse on.

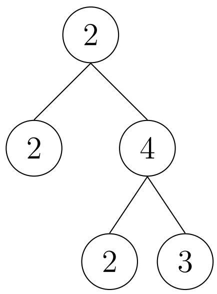
1. Write the function `even_square_tree`, which takes in a tree `t` and returns a new tree with only the even labels squared.

```
def even_square_tree(t):
    """
    >>> t = tree(2, [tree(1), tree(4)])
    >>> even_square_tree(t)
    tree(4, [tree(1), tree(16)])
    """
    ___________________________________________

    if ________________________________________:

        return ________________________________

    else:

        return ________________________________
```

```
def even_square_tree(t):
    branches_s = [even_square_tree(b) for b in branches(t)]
    if label(t) % 2 == 0:
        return tree(label(t) * label(t), branches_s)
```

```
    else:
        return tree(label(t), branches_s)
```

2. Write a function, `all_paths` that takes in a tree, `t`, and returns a list of paths from the root to each leaf. For example, if we called `all_paths(t)` on the following tree:



`all_paths(t)` would return `[[2, 2], [2, 4, 2], [2, 4, 3]]`.

```
def all_paths(t):
    paths = []

    if ________________________________________

        __________________________________________
    else:

        __________________________________________

            ______________________________________

                __________________________________
    return paths
```

```
def all_paths(t):
    paths = []
    if is_leaf(t):
        paths += [[label(t)]]
    else:
        for b in branches(t):
            for path in all_paths(b):
                paths += [[label(t)] + path]
    return paths
```

Explanation: We begin by making a list to contain all the paths.

If the tree is a leaf, the root is a leaf, so the only path is `[label(t)]`.

Otherwise, for each branch in the tree, we can use recursion to generate all the paths that extend from that branch to a leaf.

Finally, we combine the root label with each branch-starting path to make it a path from the root to a leaf.

Append every path like this to `paths`, and we have created a list of all paths!

3. Write a function that returns `True` if and only if there exists a path from root to leaf that contains at least `n` instances of `elem` in a tree `t`.

```
def contains_n(elem, n, t):
    """
    >>> t1 = tree(1, [tree(1, [tree(2)])])
    >>> contains(1, 2, t1)
    True
    >>> contains(2, 2, t1)
    False
    >>> contains(2, 1, t1)
    True
    >>> t2 = tree(1, [tree(2), tree(1, [tree(1), tree(2)])])
    >>> contains(1, 3, t2)
    True
    >>> contains(2, 2, t2) # Not on a path
    False
    """
    if n == 0:

        return True

    elif ____________________________________________:

        return ______________________________________

    elif label(t) == elem:

        return ______________________________________

    else:

        return ______________________________________
```

```python
def contains_n(elem, n, t):
    if n == 0:
        return True
    elif is_leaf(t):
        return n == 1 and label(t) == elem
    elif label(t) == elem:
        return True in [contains_n(elem, n - 1, b) for b in
          branches(t)]
    else:
        return True in [contains_n(elem, n, b) for b in
          branches(t)]
```