Homework 1
COSE312, Spring 2020

Hakjoo Oh

Due: 04/10, 23:59

Problem 1 The goal of this assignment is to write a compiler that translates regular expressions to deterministic finite automata (DFAs).

1. Clone the Git repository for programming assignments:

   ```
   git clone https://github.com/kupl/Compilers2020.git
   ```

2. You can find the `hw1` directory and the following files in it:

   - `main.ml`: Driver code with some test cases. You can add your own test cases here.
   - `regex.ml`: The definition of regular expressions (the "source language" of our compiler).
   - `nfa.ml`: NFA implementation (the "intermediate representation" of our compiler). Read `nfa.mli` to see how to use the NFA module.
   - `dfa.ml`: DFA implementation (the "target language" of our compiler). Read `dfa.mli` to see how to use the DFA module.
   - `hw1.ml`: Complete and submit this file.

3. In `regex.ml`, regular expression is defined as follows:

   ```
   type alphabet = A | B
   type t =
     | Empty
     | Epsilon
     | Alpha of alphabet
     | OR of t * t
     | CONCAT of t * t
     | STAR of t
   ```

   where we assume $\Sigma = \{\texttt{A}, \texttt{B}\}$.

4. In `hw1.ml`, you can find code below:

```
let regex2nfa : Regex.t -> Nfa.t
=fun regex -> raise Not_implemented (* TODO *)

let nfa2dfa : Nfa.t -> Dfa.t
=fun nfa -> raise Not_implemented (* TODO *)

(* Do not modify this function *)
let regex2dfa : Regex.t -> Dfa.t
=fun regex ->
  let nfa = regex2nfa regex in
  let dfa = nfa2dfa nfa in
    dfa

let run_dfa : Dfa.t -> alphabet list -> bool
=fun dfa str -> raise Not_implemented (* TODO *)
```

Your job in this assignment is to implement the functions:

- `regex2nfa`, which converts a regular expression to an equivalent NFA,
- `nfa2dfa`, which converts an NFA to an equivalent DFA, and
- `run_dfa`, which takes a DFA and a string (i.e., a sequence of input symbols) and returns true (i.e., accept) or false (i.e., reject).

Once you complete the implementation, you can build and run the program as follows:

```
$ ./build
$ ./main.native
```

For the test cases in `main.ml`:

```
let testcases : (Regex.t * alphabet list) list =
  [
    (Empty, []);
    (Epsilon, []);
    (Alpha A, [A]);
    (Alpha A, [B]);
    (OR (Alpha A, Alpha B), [B]);
    (CONCAT (STAR (Alpha A), Alpha B), [B]);
    (CONCAT (STAR (Alpha A), Alpha B), [A;B]);
    (CONCAT (STAR (Alpha A), Alpha B), [A;A;B]);
    (CONCAT (STAR (Alpha A), Alpha B), [A;B;B]);
    (CONCAT (CONCAT (STAR (CONCAT (Alpha A, Alpha A)),
                     STAR (CONCAT (Alpha B, Alpha B))), Alpha B), [B]);
    (CONCAT (CONCAT (STAR (CONCAT (Alpha A, Alpha A)),
                     STAR (CONCAT (Alpha B, Alpha B))), Alpha B), [A;A;B]);
    (CONCAT (CONCAT (STAR (CONCAT (Alpha A, Alpha A)),
```

```
                    STAR (CONCAT (Alpha B, Alpha B))), Alpha B), [B;B;B]);
    (CONCAT (CONCAT (STAR (CONCAT (Alpha A, Alpha A)),
                    STAR (CONCAT (Alpha B, Alpha B))), Alpha B), [A;A;A;A;B;B;B]);
    (CONCAT (CONCAT (STAR (CONCAT (Alpha A, Alpha A)),
                    STAR (CONCAT (Alpha B, Alpha B))), Alpha B), [A;A;A;B;B;B])
  ]
```

the correct output would be the following:

```
false
true
true
false
true
true
true
true
false
true
true
true
true
true
false
```