

2 Foundations of Computer Science (jjl25)

A certain encoding of data treats bit sequences as paths down a tree, where 0 means taking the left node, and 1 means taking the right node. When a leaf is reached, the token associated with the leaf is emitted and decoding starts again at the root for the next input digit.

Consider for example the following trees:



The bit sequence 11100 is decoded as “charlie”, “bob”, “alice” using the first tree and 2.718, 2.718, 2.718, 1.414 using the second.

Using the `seq` and `tree` types defined as follows

```
type 'a seq = Nil | Cons of 'a * (unit -> 'a seq)
type 'a tree =
  | Node of 'a tree * 'a tree
  | Leaf of 'a
```

the input bits are supplied as a `bool seq` where `true` represents 1 and `false` represents 0.

- (a) Write a function `decode : 'a tree -> bool seq -> 'a seq` that decodes the supplied bit sequence and returns a sequence of the decoded tokens. You should assume that the input stream is infinitely long. [10 marks]
- (b) Write a function `scan : ('a -> 'b -> 'b) -> 'a seq -> 'b -> 'b seq`, and show how this would be used with a suitable function and a sequence of type `float seq` to produce a running total, where each float in the output sequence is the sum of the floats seen so far in the input sequence. [6 marks]
- (c) What do the types of `decode` and `scan` tell you about the relationships between the inputs and the results of those functions? [4 marks]