Research Notebook (2022)

Boaty McBoatface

Compiled on: February 15, 2023

Table of Contents

CHAPTER 1

eResearch 2023-02-15 (Wednesday)

Stream A1

Max Wilkinson:Building a Coherent and Consistent National Research Data Storage Investment Mode

- Australian-Research Data Commons
- test note

Anton Angelo: Why do researchers say they'll share their data, and then don't?

- Best practice to publish data at the same time as manuscript

Ryan Perry: A data risk assessment and education tool for the Australian Data Archive using R Shiny

- Human focused.
- Tool used to increase anonimity for human participants in studies

Claire Rye / Greg Darcy : Accelerating data movement between New Zealand and Australia

- Globus - infrasructure to transport large data sets.
- AARNet - non-profit owned by Aus Unis and CSIRO that uses Globus.
-

Stream A3

Nisha Ghatak: Building Research Capabilities With The Carpentries

- test note

CHAPTER 2

eResearch 2023-02-16 (Thursday)

Annie West: Utilising Aotearoa's computing capability in microbial ecology: from kākāpō to estuarie

- Applications for microbial ecology for invasive species?

Appendices

A section in my appendices!

/Users/eper363/Projects/Poolean/Poolean/HuffmanDomain.fs

```fsharp
module Poolean.HuffmanDomain

open NucleotideDomain
open CompressionDomain

module HuffmanDomain =

    type ASCIICount = { Key: char; Value: int} // could change for n-grams (string) and frequency (float)

    /// 1's based Heap implementation (Priority Queue)
    /// - Smallest (integer) items have the highest priority
    ///
    /// See Kleinberg & Tardos for more details (2nd edition, p. 60)
    /// - This implementation is not thread-safe
    type Heap() =
        let mutable size = 0
        let mutable queue = [| {Key = '\000'; Value = 0} |] // first index is a placeholder, don't use it

        let swap e1 e2 =
            if (queue.[e2]).Value < (queue.[e1]).Value then
                let tmp = queue.[e1]
                queue.[e1] <- queue.[e2]
                queue.[e2] <- tmp

        member self.Print() = printfn "Queue:\t%A\nSize:\t%d" queue size

        /// find parent index of node i
        member self.Parent(i) = floor((i |> float) / 2.0) |> int

        /// bubble values up the heap
        member self.HeapifyUp(i) =
            if i > 1 then
                let j = self.Parent(i)
                swap j i
                self.HeapifyUp(j) // more efficient if I checked i < j, this will go up the tree with NOPs

        /// bubble values down the heap
        member self.HeapifyDown(i) =
            let n = size
            match (2*i > n), (2 * i < n), (2 * i = n) with
            | _, true, _ ->
                let left = 2*i
                let right = 2*i + 1
                (match (queue.[left]).Value < (queue.[right]).Value with | true -> left | false -> right) |> Some
            | _, _, true -> 2*i |> Some
            | _, _, _ -> None
            |> Option.bind (fun j -> swap i j; self.HeapifyDown(j); Some j) |> ignore

        /// insert value into heap H
        /// use heapify-up to repair damaged heap structure after each call
        /// new elements get appended to the end of the internal array
        member self.Insert(v) =
            queue <- Array.append queue [|v|]
            size <- size + 1
            self.HeapifyUp(size)

        /// if heap contains elements, return minimum element
        member self.FindMin() = match size >= 1 with | true -> Some queue.[1] | false -> None

        /// Delete element in heap position i
        /// use heapify-down to repair damaged heap structure after each call
        member self.Delete(i) =
            queue <- Array.append queue.[0..i - 1]  queue.[i+1 .. size]
            size <- size - 1
            self.HeapifyDown(i)

        /// identify and delete element with minimum key value
        member self.ExtractMin() = self.FindMin() |> Option.bind (fun min -> self.Delete(1); Some min)
```

Page 1 of 1

Figure 2.1: A priority queue is required for the Huffman coding algorithm. This is one possible implementation, from Kleinberg and Tardos, 2006.

Bibliography

Kleinberg, J., & Tardos, É. (2006). Algorithm design. Pearson.