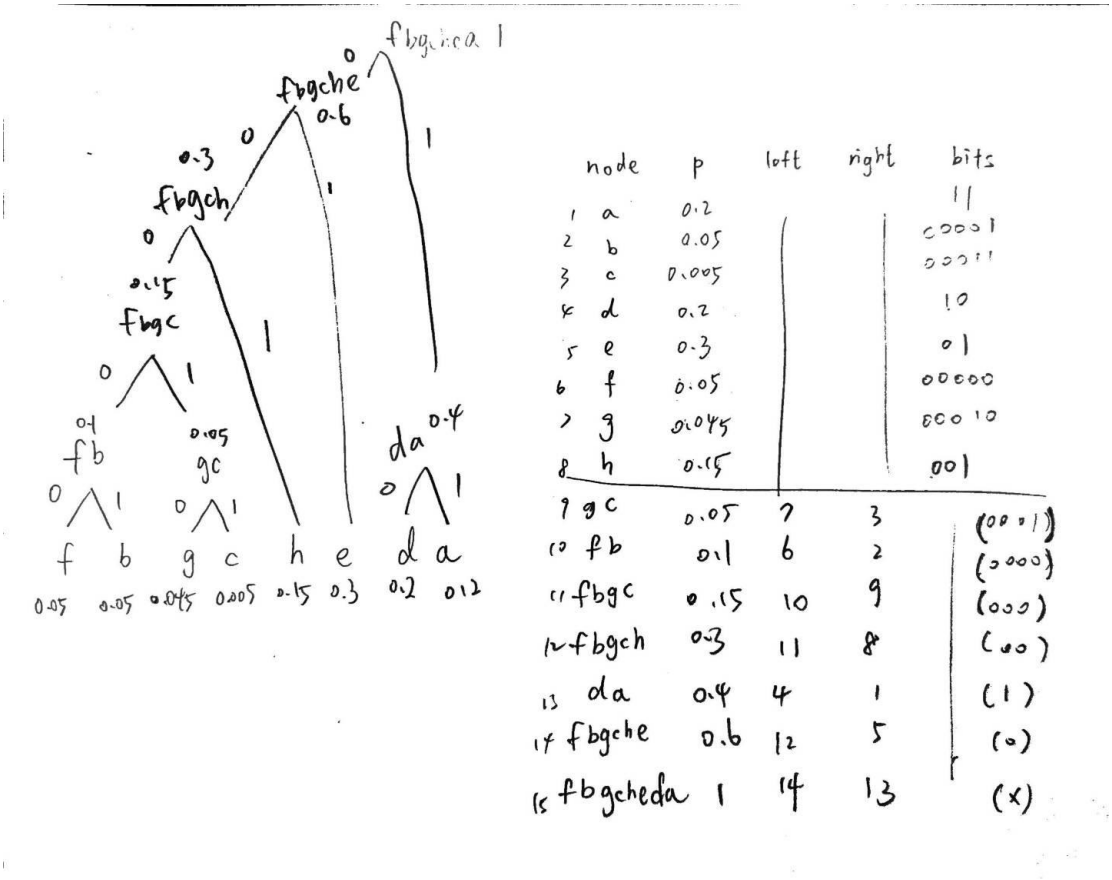
CommLab_4 Report b08901019 吳冠緯

1. Information theory and Huffman coding
   (a) Entropy = H[X] = 2.5321
   (b)



   (c) 0.25+0.03125+0.03125+0.25+0.25+0.03125+0.03125+0.125
   = 0.5 + 0.375 + 0.125 = 1 → satisfy
   (d) (2+5+5+2+2+5+5+3)/8 = 3.625 > 2.5321 → satisfy
   (e) '00010110000111100001'
   (f) 'g, a, c, a, b'
   (g)
   List 10 members
   {aaaaeeehbf}
   {aaadeeehbb}
   {aaddeeehfb}
   {adddeeehff}
   {ddddeeehbf}
   {dddaeeehbb}

{ddaaeeehfb}
{daaaeeehff}
{eeeaaaahbb}
{eeeaaadhfb}
ABS(-log2(0.027 * 0.04 * 0.04 * 0.15 * 0.0025) / 10 – 2.5321) = 0.056 < 0.1

2. Implementation of Huffman coding

(a)

```
dict =

  15×5 cell array

    {'a'       }    {[0.2000]}    {0×0 double}    {0×0 double}    {'11'      }
    {'b'       }    {[0.0500]}    {0×0 double}    {0×0 double}    {'00001'   }
    {'c'       }    {[0.0050]}    {0×0 double}    {0×0 double}    {'00011'   }
    {'d'       }    {[0.2000]}    {0×0 double}    {0×0 double}    {'10'      }
    {'e'       }    {[0.3000]}    {0×0 double}    {0×0 double}    {'01'      }
    {'f'       }    {[0.0500]}    {0×0 double}    {0×0 double}    {'00000'   }
    {'g'       }    {[0.0450]}    {0×0 double}    {0×0 double}    {'00010'   }
    {'h'       }    {[0.1500]}    {0×0 double}    {0×0 double}    {'001'     }
    {'gc'      }    {[0.0500]}    {[       7]}    {[       3]}    {'0001'    }
    {'fb'      }    {[0.1000]}    {[       6]}    {[       2]}    {'0000'    }
    {'fbgc'    }    {[0.1500]}    {[      10]}    {[       9]}    {'000'     }
    {'fbgch'   }    {[0.3000]}    {[      11]}    {[       8]}    {'00'      }
    {'da'      }    {[0.4000]}    {[       4]}    {[       1]}    {'1'       }
    {'fbgche'  }    {[0.6000]}    {[      12]}    {[       5]}    {'0'       }
    {'fbgcheda'}    {[     1]}    {[      14]}    {[      13]}    {0×0 double}
```

(b)

```
bin_seq =

    "0001011000111100001"
```

same

(c)

```
sym_seq =

  1×5 cell array

    {'g'}    {'a'}    {'c'}    {'a'}    {'b'}
```

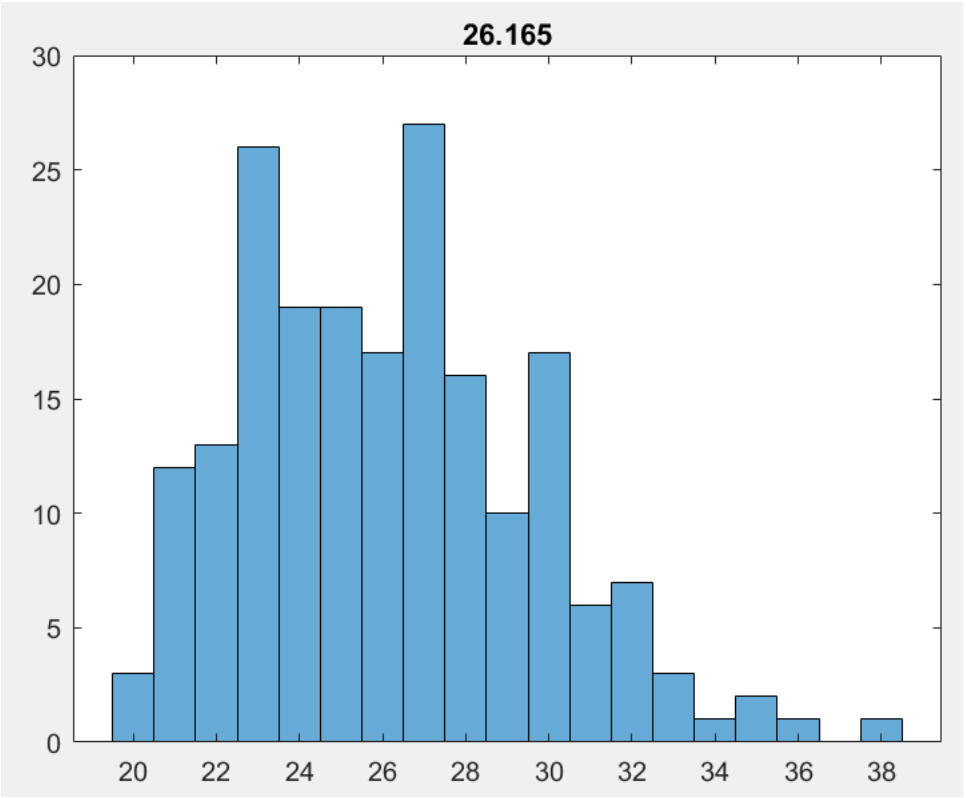
same

3. The average codeword length of Huffman coding

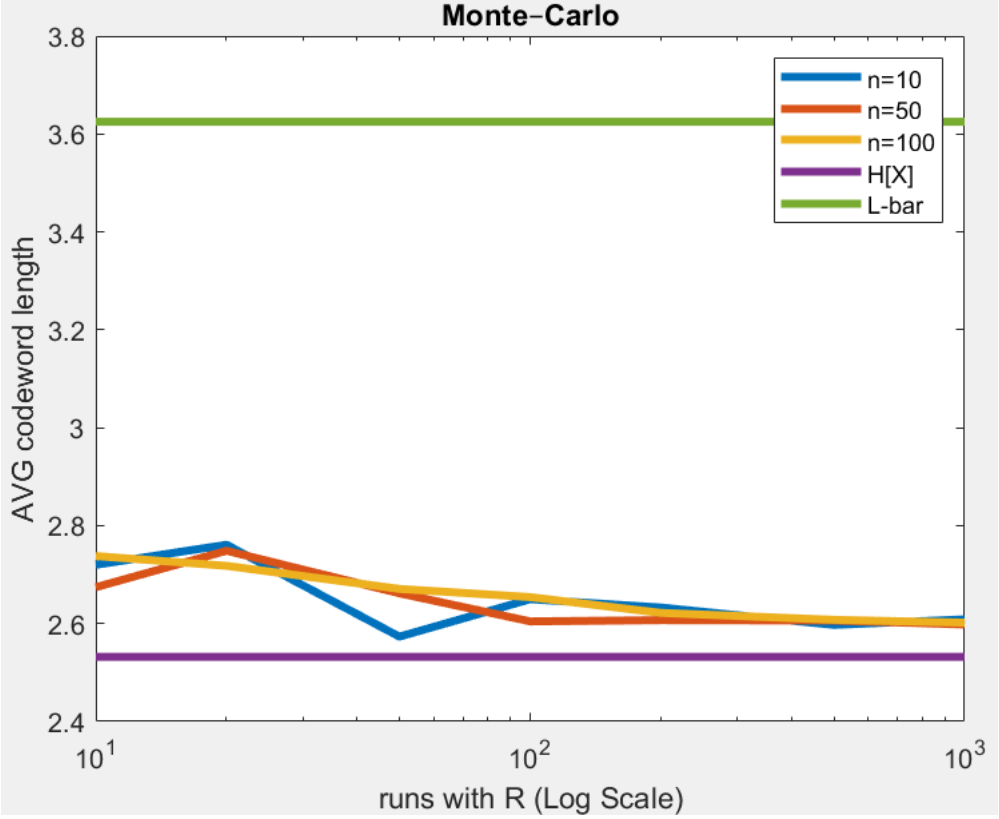
(a)

```
ten_rand =

  1×10 cell array

    {'a'}    {'a'}    {'e'}    {'e'}    {'e'}    {'g'}    {'e'}    {'b'}    {'d'}    {'h'}


ten_bin =

    "11110101010001001000011000 1"


ten_sym =

  1×10 cell array

    {'a'}    {'a'}    {'e'}    {'e'}    {'e'}    {'g'}    {'e'}    {'b'}    {'d'}    {'h'}


LEN =

    27
```

(b) Mean = 26.165



(c)



(d)

從上圖可以觀察到幾個結果，首先 Huffman code 的結果是小於 average code length 的，其次 entropy 對於收斂到 2.6 多的結果來說可以是一個下界。對於中間不同的三條曲線，若根據 R 來看，隨著 R 增大三條線都有所趨緩並收斂至同一個值，根據 n 來看可以發現 n 越大的 case 在模擬過程中變異程度越小，意即浮動程度越不明顯。

心得

這次實驗是時隔一年多再次使用 matlab 進行實驗模擬，有許多語法已經忘記，花了不少額外的時間。不過 Huffman code 的實做在之前演算法的課程中已有所認識，再加上這次上課又

聽了一次，整體上花費的時間還是差不多的。