

アルゴリズムとデータ構造III 15回目

期末試験

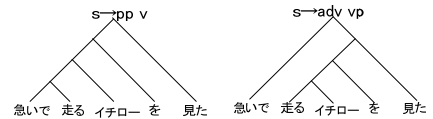
問題1 構文解析CKY法

- 下の図は「急いで走るイチローを見た」を構文解析中のCKY表である。
- 図中の①, ②, ③, ④, ⑤には何が入るか答えよ。
- CKY表から得られる「急いで走るイチローを見た」の構文木を描け。

| | | | | | |
|------|----|------|----|----|-----|
| 急いで | 走る | イチロー | を | 見た | |
| 急いで | 走る | イチロー | を | 見た | ② ⑤ |
| 走る | 走る | pp | v | n | ① ④ |
| イチロー | s | adv | pp | n | ③ |
| を | | p | を | | |
| 見た | | v | 見た | | |

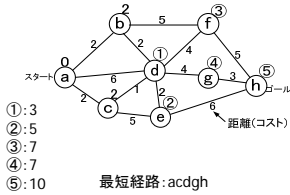
- 書き換え規則
- (1) s → pp v
 - (2) s → adv vp
 - (3) vp → pp v
 - (4) vp → adv v
 - (5) np → vp n
 - (6) np → v n
 - (7) pp → np p
 - (8) pp → sp p
 - (9) adv → 急いで
 - (10) n → イチロー
 - (11) p → を
 - (12) v → 走る
 - (13) v → 見た

- ①: (7)
- ②: (7)
- ③: (1), (3)
- ④: (1), (3)
- ⑤: (1), (2), (3)



問題2 ダイクストラ法

- 下の図はダイクストラ法による最短経路探索を行っているグラフを表している。図中の①~⑤はスタートノードから各ノードへの最短距離である。①~⑤にはそれぞれ何が入るか答えよ。
- またスタートノードからゴールノードまでの最短経路を答えよ。



問題3 DPマッチング

- 下の表は「abcd」と「bccd」の単語間距離をDPマッチングにより計算しているところである。表中の①, ②, ③, ④, ⑤には何が入るか答えよ。但し、不一致ペナルティは3点、1字ずれた場合のペナルティは1点とする。

通行ペナルティ積算表

| | | | | | |
|---|----|----|---|----|--|
| | a | b | c | d | |
| b | 3 | 4 | 8 | 12 | |
| c | 7 | 6 | 4 | ① | |
| c | 11 | 10 | ② | ③ | |
| d | 15 | 14 | ④ | ⑤ | |

- ①: 8
- ②: 5
- ③: 7
- ④: 9
- ⑤: 5

問題4 KMP法

- 下の表はKMP法のNEXT関数値を表している。表中の①~⑤には何が入るか答えよ。但し、NEXT関数とは「文字照合に失敗したときに次回の照合でキーワードの何文字目を照合すべきか」を表す関数である。

| | | | | | | | |
|---------|---|---|---|---|---|---|---|
| 位置 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| キーワード | a | b | c | a | b | a | |
| NEXT関数値 | 0 | 1 | ① | ② | ③ | ④ | ⑤ |

- ①: 1
- ②: 0
- ③: 1
- ④: 3
- ⑤: 2

問題5 BM法

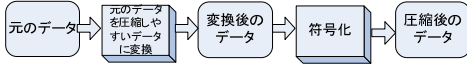
- 下の図はキーワード「abcaba」に対するBM法のskip関数値を表している。表中の①~④には何が入るか答えよ。但し、skip関数とは文字照合に失敗した場合、キーワードを何文字右にシフトして次回の照合を行えば良いかを表す。

| | |
|---------|---------|
| 文字 | Skip関数値 |
| a | ① |
| b | ② |
| c | ③ |
| 上記以外の文字 | ④ |

- ①: 2
- ②: 1
- ③: 3
- ④: 6

問題6 データ圧縮

- 下の図は一般的なデータ圧縮処理の流れを示している。「圧縮しやすいデータ」の特徴を2つ挙げよ。



出現する事象の確率の偏りが大きい
出現する事象の数が少ない

問題7 ハフマン符号化1

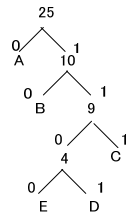
- ハフマン符号について簡単に説明せよ。

2分木を使って頻度の高い記号に短い符号を、頻度の低い記号に長い符号を割り当てる左から符号列をトレースすることで符号の区切りを一意に決めることができる。

問題8 ハフマン符号化2

- 下の表のような記号の出現頻度のとき、ハフマン符号をつくりなさい。但しハフマン符号作成のための二分木も書くこと。

| 記号 | 頻度 | Huffman code |
|----|----|--------------|
| A | 9 | 0 |
| B | 7 | 10 |
| C | 5 | 111 |
| D | 3 | 1101 |
| E | 1 | 1100 |



問題9 音声データの圧縮

- 多くの音声データ圧縮手法では、音声波形データを周波数表現に変換する。周波数表現に変換する利点を説明せよ。
- 人間の耳で聞こえないような高周波数領域をカットすることでデータ量を削減できる。
- サンプリング周期を長くすることでデータ量を削減できる。
- ベクトル量子化、線形予測を利用できる。

問題10 DCT

- DCTを簡単に説明せよ。
- 離散コサイン変換
- JPEGなどの画像圧縮に使われる。
- 8x8ピクセルのブロックに分割し、ブロックごとに2次元の周波数領域に変換する。
- 人間の視覚特性に合わせて高周波成分は粗く量子化し、低周波成分は細かく量子化することで圧縮効率を高めることができる。