

アルゴリズムとデータ構造III 15回目

期末試験

問題1 問題1. スタック, 記法

(平成22年度春期基本情報技術者 午前 問03)

- A=1, B=3, C=5, D=4, E=2のとき, 逆ポーランド表記法で表現された式 $AB+CDE/-*$ の演算結果はどれか.
- 但し, $AB/ \rightarrow A \div B$, $AB- \rightarrow A-B$
- ア -12, イ 2, ウ 12, エ 14
- ウ 12

1	3	+	4	5	4	2	7	2	5	3	*	12
	1	3		4	4	4	2	5	2	4	3	
		1		4	4	5	4	4	5	4	3	4
				4	4	4	4	4	5	4	3	4
				4	4	4	4	4	4	4	3	4

問題2 DPマッチング

- 下の表は「abcde」と「abbde」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①～⑤には何が入るか答えよ。但し、不一致ペナルティは3点、脱落ペナルティは1点、挿入ペナルティは1点とする。

通行ペナルティ積算表

	a	b	c	d	e
a	0	4	8	12	16
b	4	0	4	8	12
b	8	1	3	7	11
d	12	5	4	①	②
e	16	9	③	④	⑤

- ① 3
- ② 7
- ③ 8
- ④ 7
- ⑤ 3

問題3 文字列検索 (KMP法)

- 下の表はキーワード「abcaba」に対するKMP法のNEXT関数値を示している。表中の①～⑤には何が入るか答えよ。但し、NEXT関数とは「文字列照合に失敗したときに次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4 文字列検索 (BM法)

- 下の表はキーワード「abcaba」に対するBM法のskip関数値を示している。表中の①～④には何が入るか答えよ。但し、skip関数とは「文字列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の照合を行えばよいか」を表す関数である。

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

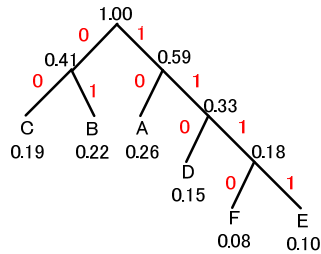
問題5 文字列検索 (Aho-Corasick法)

- Aho-Corasick法を100文字程度で説明せよ。
- 複数のキーワードを同時に検索できる文字列検索アルゴリズム
- KMP法を複数キーワードに適用した方法
- 複数キーワードより有限オートマトンを作成し、文字列を入力とすることで文字列照合を行う。
- バックトラックを行わずに複数キーワードを同時に検索できる

問題6 ハフマン符号化

- 下の表のような記号の出現確率のとき、ハフマン符号を作りなさい。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	符号
A	0.26	10
B	0.22	01
C	0.19	00
D	0.15	110
E	0.10	1111
F	0.08	1110



問題7 エントロピー

- アルファベットのAからDまでの4文字が全て等確率で出現すると仮定した場合の平均情報量(エントロピー)求めよ。また各文字の出現確率が下の表である場合、平均情報量(エントロピー)を求めよ。

文字	A	B	C	D
出現確率	1/2	1/4	1/8	1/8

$$\text{等確率: } H = -4 \times \frac{1}{4} \log_2 \frac{1}{4} = -\log_2 2^{-2} = 2 \text{ ビット(シャノン)}$$

$$\begin{aligned} \text{表確率: } H &= -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{8} \log_2 \frac{1}{8} - \frac{1}{8} \log_2 \frac{1}{8} \\ &= -\frac{1}{2} \log_2 2^{-1} - \frac{1}{4} \log_2 2^{-2} - \frac{1}{8} \log_2 2^{-3} - \frac{1}{8} \log_2 2^{-3} \\ &= \frac{1}{2} + \frac{1}{4} + 2 \times \frac{3}{8} = \frac{7}{4} \text{ ビット(シャノン)} \end{aligned}$$

問題8 音声圧縮

- ADPCMを100字程度で説明せよ。
- Adaptive Differential Pulse Code Modulation
- 適応的差分パルス符号変調
- 音声波形は連続的に変化しているので前回のサンプリングからの差分を記録するだけなら量子化ビット数を抑えられる
- 大きな変化の部分では量子化幅を荒くすることで効率よく圧縮できる

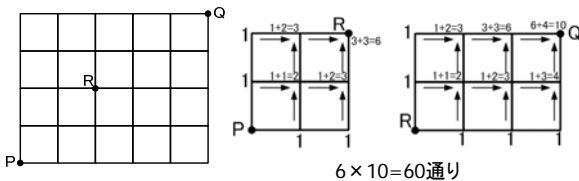
問題9 DCT

- 画像圧縮に用いられるDCTについて100文字程度で説明せよ。
- 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。

問題10 最短経路の組み合わせ

(平成20年秋期基本情報技術者午前問07改)

- 図の線上を点Pから点Rを通過して、点Qに至る最短経路は何通りあるか。ここで、縦1区画の長さはすべて等しく、横1区画の長さもすべて等しいものとする。



$$\binom{4}{2,2} \times \binom{5}{2,3} = \frac{4!}{2!2!} \times \frac{5!}{2!3!} = 6 \times 10 = 60 \text{ 通り}$$