

情報工学

予想問題 第6回 解答

(基礎技法の総点検年)

記述式の小問については解答例を示す。同じ内容が伝われば表現は問わない。

作成：Claude (Anthropic)

1 【必須問題】 アルゴリズムとプログラミング(1-1) (a) $(i + 1) \% M$ (b) $(i + 1) \% M$

線形探索法では衝突が起きたとき次の番地を順に調べる。表の末尾に達したら先頭に戻る必要があるため、 M による剰余をとる。

(1-2) $h(k) = k \bmod 11$ により順に挿入する。

キー	$h(k)$	挿入位置	ループ実行回数
22	0	0	0
15	4	4	0
8	8	8	0
33	0	1 (0 が使用中)	1
4	4	5 (4 が使用中)	1
26	4	6 (4,5 が使用中)	2

完成したハッシュ表は以下のとおり。

添字	0	1	2	3	4	5	6	7	8	9	10
値	22	33			15	4	26		8		

(2-1) $\text{search}(26)$: $h(26) = 4$ から順に $T[4] = 15$, $T[5] = 4$, $T[6] = 26$ を調べる。3 回目で一致し添字 6 を返す。このとき count は 3 増える。

$\text{search}(8)$: $h(8) = 8$ で $T[8] = 8$ が直ちに一致し添字 8 を返す。 count は 1 増える。

したがって count の最終値は $3 + 1 = 4$ であり、出力は以下のとおり。

6
8
4

(2-2) $h(37) = 37 \bmod 11 = 4$ である。添字 4, 5, 6, 7 の順に参照し、 $T[7]$ が EMPTY であるため探索は失敗する。

参照される要素の添字は 4, 5, 6, 7 の順であり、 count が増える回数は 4 回である（ループ内で 3 回、ループを抜けた後の 27 行目で 1 回）。

(3-1) $n = 6$, $M = 11$ であるから

$$\alpha = \frac{6}{11} \approx 0.545.$$

成功時の平均比較回数は

$$\frac{1}{2} \left(1 + \frac{1}{1 - 0.545} \right) = \frac{1}{2} (1 + 2.20) \approx \mathbf{1.60}.$$

失敗時の平均比較回数は

$$\frac{1}{2} \left(1 + \frac{1}{(1 - 0.545)^2} \right) = \frac{1}{2} (1 + 4.84) \approx \mathbf{2.92}.$$

$\alpha \rightarrow 1$ のとき： $1/(1 - \alpha)$ および $1/(1 - \alpha)^2$ が発散するため、平均比較回数が急激に増大する。表がほぼ満杯になると、空き番地を見つけるまでに表の大部分を走査することになり、ハッシュ法の利点である $O(1)$ に近い性能が失われる。実用上は α を 0.7 程度以下に保ち、超えたら表を拡張して再ハッシュ (rehash) する。

(3-2) 現象の名称：一次クラスタリング (primary clustering)。

緩和する探査方法の例：二次探査法 (quadratic probing)。 i 回目の探査位置を

$$h_i(k) = (h(k) + c_1 i + c_2 i^2) \bmod M \quad (c_2 \neq 0)$$

とする。探査の刻み幅が i とともに変化するため、同じハッシュ値をもつキーどうしが連続した塊を作りにくくなる。

(別解として二重ハッシュ法 (double hashing) $h_i(k) = (h_1(k) + i h_2(k)) \bmod M$ も可。)

(3-3) 具体例：小問 (1-2) の表からキー 4 (添字 5) を削除して $T[5]$ を EMPTY に戻したとする。この状態でキー 26 を探索すると、 $h(26) = 4$ から $T[4] = 15$ (不一致)、次に $T[5]$ を見るがここが EMPTY であるためループを抜け、-1 (見つからない) を返してしまう。実際には 26 は添字 6 に存在するにもかかわらず発見できない。

原因は、削除によって探査の連鎖が途中で切断されることである。

回避方法:削除した位置に「かつて使用されていた」ことを表す特別な標識(**墓標, tombstone**)を書き込む。探索時はこの標識を空きとみなさず探査を継続し、挿入時は空き番地として再利用できるようにする。

2 【必須問題】 計算機システムとシステムプログラム

(1-1) $45 = 00101101_2$ である。2 の補数は「全ビット反転して 1 を加える」ことで得られる。

$$00101101 \xrightarrow{\text{反転}} 11010010 \xrightarrow{+1} 11010011.$$

よって

$$-45 = 11010011_2 = 0xD3.$$

(検算： 11010011_2 を符号付きとして解釈すると $-128 + 64 + 16 + 2 + 1 = -45$ となり一致する。)

(1-2) 表せる範囲は

$$-128 \leq x \leq 127 \quad (-2^7 \leq x \leq 2^7 - 1).$$

一致しない理由：0 が正の側 (00000000) に 1 つ含まれるため、0 を除く正の数は 127 個であるのに対し負の数は 128 個となる。すなわち 0 の分だけ正の側が 1 つ少ない。

(1-3) $100 = 01100100_2$, $50 = 00110010_2$ であるから

$$01100100 + 00110010 = 10010110_2.$$

これを符号付き 8 ビット整数として解釈すると最上位ビットが 1 であるから負の数であり、

$$10010110_2 \rightarrow -(01101010_2) = -106.$$

すなわち $100 + 50 = 150$ となるべきところ -106 という誤った結果になる。

検出方法：二つの被加数の最上位ビットが等しく、かつ加算結果の最上位ビットがそれと異なる場合にオーバーフロー (overflow) が発生している。本問では被加数がともに 0 (正) であるのに結果が 1 (負) となっているため検出できる。(符号が異なる数どうしの加算ではオーバーフローは生じない。)

(1-4) $x = -128 = 10000000_2$ である。全ビットを反転すると 01111111 , これに 1 を加えると

$$01111111 + 1 = 10000000_2$$

となり、元の -128 に戻ってしまう。

正しくない理由： $+128$ は 8 ビットの 2 の補数表現で表せる範囲 (最大 127) を超えているため、そもそも表現できない。小問 (1-2) で述べたとおり負の数が正の数より 1 つ多いという非対称性があり、最小値 -128 だけは符号反転がこの表現の中で閉じない。

(2-1) **長所**：ファイルの先頭ブロック番号と長さだけで位置が定まるため、任意の位置へのアクセス (直接アクセス, random access) が計算のみで可能であり高速である。またヘッドの移動が少なく順次読み出しの性能も高い。

短所：ファイルの作成時に必要な大きさの連続領域を確保する必要があり、ファイルの拡張が困難である。また使用と解放を繰り返すうちに小さな空き領域が散在し、外部断片化 (external fragmentation) が生じる。

(2-2) 1 個の間接ブロックに格納できるポインタの個数は

$$\frac{4 \text{ KiB}}{4 \text{ バイト}} = \frac{4096}{4} = 1024 \text{ 個}.$$

表現できるブロック数は

$$12 + 1024 + 1024^2 + 1024^3 = 12 + 1024 + 1048576 + 1073741824 = 1074791436 \text{ ブロック}.$$

最大ファイルサイズは

$$1074791436 \times 4096 \text{ バイト} = 4402345721856 \text{ バイト} \approx \mathbf{4.00 \text{ TiB.}}$$

(実質的には 3 段間接の 1024^3 ブロック = 4 TiB がほぼすべてを占める。)

(2-3) 10 MiB = 10×2^{20} バイトであるから、必要なブロック数は

$$\frac{10 \times 2^{20}}{4096} = \frac{10485760}{4096} = \mathbf{2560 \text{ ブロック.}}$$

内訳は以下のとおり。直接ポインタで参照できるのは 12 個、1 段間接ポインタで参照できるのは 1024 個であり、

$$12 + 1024 = 1036 < 2560$$

であるから、これらだけでは足りず 2 段間接ポインタも用いる必要がある。

したがって

$$\text{直接 : } \mathbf{12 \text{ 個}}, \quad \text{1 段間接 : } \mathbf{1024 \text{ 個}}, \quad \text{2 段間接 : } 2560 - 1036 = \mathbf{1524 \text{ 個}}$$

となる。

(2-4) 先頭から 10 MiB 目付近は最終ブロック、すなわち 2 段間接ポインタが指す領域にある。i-node が主記憶上にあるとすると、

1. 2 段間接ブロック (1 段目) を読む.....1 回
2. そこが指す間接ブロック (2 段目) を読む.....1 回
3. 目的のデータブロックを読む.....1 回

で合計 **3 回** のディスクアクセスが必要である。

連結割り当ての場合：先頭ブロックから順にポインタをたどる必要があるため、2560 番目のブロックに到達するには **2560 回** のディスクアクセスを要する。

利点：索引割り当てでは、ファイル内の任意の位置へのアクセスに要するディスクアクセス回数が、ファイルの大きさによらず高々 3~4 回に抑えられる。連結割り当てのように位置に比例して増加することがないため、大きなファイルへのランダムアクセスが現実的な時間で行える。

4 【選択問題】 計算理論

(1-1) $r = (a | b)^* ab$ が表す言語は、 ab で終わる語全体である。

長さ 3 以下の語で r に属するものは

$$ab, \quad aab, \quad bab$$

の 3 個である。(長さ 2 は ab のみ、長さ 3 は先頭 1 文字が a か b の 2 通り。長さ 0, 1 の語は ab で終わり得ない。)

(1-2) q_0 からは ε 遷移で q_1 に到達でき、 q_1 からはさらに ε 遷移がないので

$$\text{ECLOSE}(q_0) = \{q_0, q_1\}.$$

$\{q_0\}$ から入力 a を 1 文字読んだ後： $\text{ECLOSE}(q_0) = \{q_0, q_1\}$ の各状態から a で遷移すると、 $q_0 \xrightarrow{a} q_0$, $q_1 \xrightarrow{a} q_2$ であるから $\{q_0, q_2\}$ に到達する。 ε 閉包をとると q_0 から q_1 に行けるので

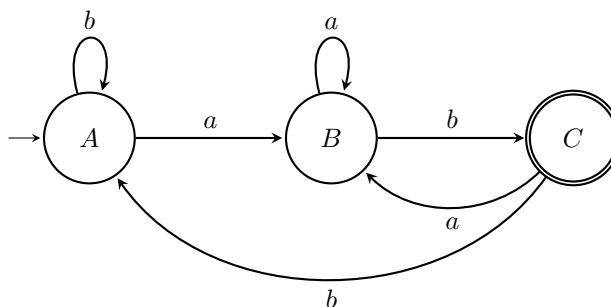
$$\text{ECLOSE}(\{q_0, q_2\}) = \{q_0, q_1, q_2\}.$$

(1-3) サブセット構成法を適用する。開始状態は $A = \text{ECLOSE}(q_0) = \{q_0, q_1\}$ である。

名前	状態集合	a	b
A	$\{q_0, q_1\}$	B	A
B	$\{q_0, q_1, q_2\}$	B	C
C	$\{q_0, q_1, q_3\}$	B	A

(B から $b : q_0 \xrightarrow{b} q_0$, $q_2 \xrightarrow{b} q_3$ より $\{q_0, q_3\}$, ε 閉包で $\{q_0, q_1, q_3\} = C$ 。 C から $a : q_0 \xrightarrow{a} q_0$, $q_1 \xrightarrow{a} q_2$ より B 。)

受理状態は q_3 を含む C のみである。状態遷移図は次のとおり。



状態数は 3。

(2-1) L_1 (a の個数が 3 の倍数) を認識する最小 DFA は 3 状態である。状態 p_0, p_1, p_2 を「これまでに読んだ a の個数を 3 で割った余り」とし、 b では状態を変えない。開始かつ受理状態は p_0 である。

$$p_0 \xrightarrow{a} p_1 \xrightarrow{a} p_2 \xrightarrow{a} p_0, \quad p_i \xrightarrow{b} p_i \quad (i = 0, 1, 2).$$

L_2 (bb を含まない) を認識する最小 DFA は 3 状態である。状態 s_0 (直前が b でない), s_1 (直前が b), s_2 (bb が出現済みの死状態) とする。受理状態は s_0, s_1 である。

$$s_0 \xrightarrow{a} s_0, \quad s_0 \xrightarrow{b} s_1, \quad s_1 \xrightarrow{a} s_0, \quad s_1 \xrightarrow{b} s_2, \quad s_2 \xrightarrow{a,b} s_2.$$

(2-2) 直積構成では状態集合が $Q_1 \times Q_2$ となるので、状態数は

$$3 \times 3 = 9.$$

$L_1 \cap L_2$ に属する空列でない最短の語： a の個数が 3 の倍数 (0 個でもよい) で、かつ bb を含まない語のうち最短のものは、 a を 0 個として b を 1 個だけ用いた $\mathbf{b}$ (長さ 1) である。

(3) $L_3 = \{a^{n^2} \mid n \geq 0\}$ が正則言語であると仮定する。反復補題により正の整数 p が存在する。

$w = a^{p^2}$ とすると $w \in L_3$ であり $|w| = p^2 \geq p$ である。よって $w = xyz$ ($|xy| \leq p$, $|y| \geq 1$) と分解でき、任意の $k \geq 0$ に対して $xy^kz \in L_3$ が成り立つ。

$y = a^m$ とおくと $1 \leq m \leq p$ である。ここで $k = 2$ をとると

$$|xy^2z| = p^2 + m.$$

$1 \leq m \leq p$ であるから

$$p^2 < p^2 + m \leq p^2 + p < p^2 + 2p + 1 = (p+1)^2.$$

すなわち $|xy^2z|$ は連続する二つの平方数 p^2 と $(p+1)^2$ の間に真に挟まれており、平方数ではない。したがって $xy^2z \notin L_3$ となり、反復補題の主張に矛盾する。

よって仮定が誤りであり、 L_3 は正則言語ではない。☒

6 【選択問題】 電子回路と論理設計

(1-1) 閉路に沿ってキルヒホッフの電圧則を適用すると

$$E = Ri(t) + v_C(t).$$

コンデンサの電流と電圧の関係 $i(t) = C \frac{dv_C}{dt}$ を代入して

$$RC \frac{dv_C(t)}{dt} + v_C(t) = E.$$

(1-2) これは1階線形微分方程式である。定常解は $v_C = E$ 、過渡解は $Ae^{-t/(RC)}$ であるから一般解は

$$v_C(t) = E + Ae^{-t/(RC)}.$$

初期条件 $v_C(0) = 0$ より $A = -E$ 。よって

$$v_C(t) = E \left(1 - e^{-t/(RC)} \right).$$

電流は

$$i(t) = C \frac{dv_C}{dt} = C \cdot E \cdot \frac{1}{RC} e^{-t/(RC)} = \frac{E}{R} e^{-t/(RC)}.$$

(1-3) 時定数は

$$\tau = RC = 2.0 \times 10^3 \times 5.0 \times 10^{-6} = 1.0 \times 10^{-2} \text{ s} = \mathbf{10.0 \text{ ms}}.$$

$t = \tau$ のとき

$$v_C = E(1 - e^{-1}) = 10 \times 0.6321 = \mathbf{6.32 \text{ V}}.$$

$t = 3\tau$ のとき

$$v_C = E(1 - e^{-3}) = 10 \times 0.9502 = \mathbf{9.50 \text{ V}}.$$

(1-4) $v_C = 0.99E$ となる条件は

$$1 - e^{-t/\tau} = 0.99 \iff e^{-t/\tau} = 0.01 \iff \frac{t}{\tau} = \ln 100 = 4.605.$$

すなわち $t \approx 4.61\tau$ である。具体的な値は

$$t = 4.605 \times 10.0 = \mathbf{46.1 \text{ ms}}.$$

(2-1) 右方向シフトでは、各段の入力に一つ左（上位）の段の出力を与える。最上位段 Y_3 には直列入力 x を与える。

$$D_3 = x, \quad D_2 = y_3, \quad D_1 = y_2, \quad D_0 = y_1.$$

初期状態 $(0, 0, 0, 0)$ から $x = 1, 0, 1, 1$ を順に与えると

クロック	入力 x	状態 (y_3, y_2, y_1, y_0)
初期	—	0, 0, 0, 0
1	1	1, 0, 0, 0
2	0	0, 1, 0, 0
3	1	1, 0, 1, 0
4	1	1, 1, 0, 1

(2-2) $D_3 = y_0 \oplus y_1$, $D_2 = y_3$, $D_1 = y_2$, $D_0 = y_1$ に従って追跡する。

0001 $\rightarrow$ 1000 $\rightarrow$ 0100 $\rightarrow$ 0010 $\rightarrow$ 1001 $\rightarrow$ 1100 $\rightarrow$ 0110 $\rightarrow$ 1011

$\rightarrow$ 0101 $\rightarrow$ 1010 $\rightarrow$ 1101 $\rightarrow$ 1110 $\rightarrow$ 1111 $\rightarrow$ 0111 $\rightarrow$ 0011 $\rightarrow$ 0001

初期状態に戻るまでに 15 回の遷移を要するので、周期は 15 である。

(0000 を除く $2^4 - 1 = 15$ 個の状態をすべて巡回しており、最大長系列 (maximum length sequence) となっている。)

(2-3) 状態 $(0, 0, 0, 0)$ を初期状態とすると、 $D_3 = 0 \oplus 0 = 0$, $D_2 = D_1 = D_0 = 0$ となるため、次状態も $(0, 0, 0, 0)$ である。すなわち**全零状態から抜け出せず、永久にその状態に留まる**。

n 個のフリップフロップを用いた線形帰還シフトレジスタでは、帰還が排他的論理和のみで構成されるため全零状態は必ず自分自身に遷移する。したがって巡回に含まれ得ない状態は全零状態の **1 個** である。

よって取り得る状態 2^n 個のうち全零状態を除いた

$$2^n - 1$$

が周期の最大値である。

7 【選択問題】 数学解析と信号処理

(1-1) 定義より

$$X(\omega) = \int_{-T/2}^{T/2} 1 \cdot e^{-j\omega t} dt = \left[\frac{e^{-j\omega t}}{-j\omega} \right]_{-T/2}^{T/2} = \frac{e^{-j\omega T/2} - e^{j\omega T/2}}{-j\omega}.$$

オイラーの公式 $e^{j\theta} - e^{-j\theta} = 2j \sin \theta$ を用いると

$$X(\omega) = \frac{2j \sin(\omega T/2)}{j\omega} = \frac{2 \sin(\omega T/2)}{\omega}.$$

これは $T \operatorname{sinc}$ の形に書き直せて

$$X(\omega) = T \frac{\sin(\omega T/2)}{\omega T/2}.$$

(1-2) $\omega \rightarrow 0$ では $\frac{\sin x}{x} \rightarrow 1$ であるから

$$\lim_{\omega \rightarrow 0} X(\omega) = T.$$

(これはパルスの面積 $\int x(t)dt = T$ に一致する。) $X(\omega) = 0$ となるのは $\sin(\omega T/2) = 0$ かつ $\omega \neq 0$ のときであるから

$$\frac{\omega T}{2} = k\pi \iff \omega = \frac{2\pi k}{T} \quad (k = 1, 2, 3, \dots).$$

最初の零点は $\omega = \frac{2\pi}{T}$ であり、パルス幅 T に反比例する。(1-3) T を小さくすると最初の零点 $2\pi/T$ が大きくなるので、 $X(\omega)$ の主要な成分が広い周波数範囲に広がる。逆に T を大きくすると $X(\omega)$ は狭い範囲に集中する。この性質は、時間領域での広がり と 周波数領域での広がり が同時には小さくできないという原理として知られており、**時間・周波数の不確定性原理** (uncertainty principle) と呼ばれる。(時間帯域幅積が一定以上になる、という表現でもよい。)

(1-4) 畳み込み定理より

$$Y(\omega) = X(\omega)^2 = \left(T \frac{\sin(\omega T/2)}{\omega T/2} \right)^2 = T^2 \left(\frac{\sin(\omega T/2)}{\omega T/2} \right)^2.$$

 $y(t)$ の波形：幅 T の矩形パルス同士の畳み込みであるから、 $y(t)$ は**三角パルス** (triangular pulse) となる。具体的には $t=0$ で最大値 T をとり、 $|t|=T$ で 0 となる二等辺三角形であり、**底辺の幅は $2T$** である。($Y(\omega) \geq 0$ となっていることも、 $y(t)$ が偶関数かつ三角形であることと整合する。)

(2-1) 畳み込み和を計算する。

$$y[0] = x[0]h[0] = 1,$$

$$y[1] = x[0]h[1] + x[1]h[0] = 1 + 2 = 3,$$

$$y[2] = x[1]h[1] + x[2]h[0] = 2 + 3 = 5,$$

$$y[3] = x[2]h[1] = 3.$$

その他の n では $y[n] = 0$ 。すなわち

$$y[n] = \{1, 3, 5, 3\} \quad (n = 0, 1, 2, 3).$$

(2-2) 定義より

$$X(z) = 1 + 2z^{-1} + 3z^{-2}, \quad H(z) = 1 + z^{-1}.$$

積を計算すると

$$\begin{aligned} Y(z) &= (1 + 2z^{-1} + 3z^{-2})(1 + z^{-1}) = 1 + z^{-1} + 2z^{-1} + 2z^{-2} + 3z^{-2} + 3z^{-3} \\ &= 1 + 3z^{-1} + 5z^{-2} + 3z^{-3}. \end{aligned}$$

z^{-n} の係数が $y[n]$ に対応するので $y[n] = \{1, 3, 5, 3\}$ となり, 小問 (2-1) の結果と一致する。

(2-3) 長さ N_1 の信号の z 変換は z^{-1} について次数 $N_1 - 1$ の多項式, 長さ N_2 の信号の z 変換は次数 $N_2 - 1$ の多項式である。畳み込みは z 変換の積に対応し, 多項式の積の次数は次数の和 $(N_1 - 1) + (N_2 - 1) = N_1 + N_2 - 2$ となる。次数 $N_1 + N_2 - 2$ の多項式の係数は 0 次から数えて $N_1 + N_2 - 1$ 個あるので, 畳み込みの長さは $N_1 + N_2 - 1$ である。

連続時間の場合との対応：小問 (1-4) では幅 T の矩形パルス同士の畳み込みが底辺の幅 $2T = T + T$ の三角パルスとなった。これは「畳み込みによって台 (support) の幅が二つの台の幅の和になる」という同じ現象であり, 離散時間の $N_1 + N_2 - 1$ (両端の重なりの分だけ 1 を引く) に対応している。