データ構造とアルゴリズムII

第4回

動的計画法 3

演習：ロッド切り出し問題

- 長さ10mの金属棒がある．これを切り出して販売する(切り出す長さは1mの倍数)．長さ i の加工された金属棒の販売価格 p_i は以下で与えられる．販売価格の合計を最大にする切り出し方法を求めよ．

length i	1	2	3	4	5	6	7	8	9	10
p_i	1	5	8	9	10	17	17	20	22	23

演習：ロッド切り出し問題

length i	1	2	3	4	5	6	7	8	9	10
p_i	1	5	8	9	10	17	17	20	22	23
v_i:	1	5	8	10	13	17	18	22	25	27

10
4-6
2-2-6

演習：ロッド切り出し問題

- ロッド切り出し問題で，切る回数が最大3等の制限を加えると，部分問題最適性が保証されなくなる
- この課題をクリアして，動的計画法を使うにはどうしたらよいか?

例題

最長共通部分列

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める
  - 部分系列とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$の部分系列が $Z=\langle z_1,z_2,\ldots,z_k\rangle$であるとは，増加する添え字の列$\langle i_1,i_2,\ldots i_k\rangle$が存在し，$x_{i_j}=z_j$ を満たす．

$X=\langle A, B, C, B, D, A, B\rangle$

部分系列

$Z=\langle B, C, D, B\rangle$

186

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める
  - 部分系列とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$の部分系列が $Z=\langle z_1,z_2,\ldots,z_k\rangle$であるとは，増加する添え字の列$\langle i_1,i_2,\ldots i_k\rangle$が存在し，$x_{i_j}=z_j$ を満たす．

$\langle 2,\ 3,\ 5,\ 7\rangle$

$X=\langle A, B, C, B, D, A, B\rangle$

部分系列

$Z=\langle B, C, D, B\rangle$

187

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める
  - 部分系列とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$の部分系列が $Z=\langle z_1,z_2,\ldots,z_k\rangle$であるとは，増加する添え字の列$\langle i_1,i_2,\ldots i_k\rangle$が存在し，$x_{i_j}=z_j$ を満たす．
  - 共通部分系列とは
    2つの系列の共通する部分系列

$X=\langle A, B, C, B, D, A, B\rangle$

$Y=\langle B, D, C, A, B, A\rangle$

$\langle B, C, A\rangle$ は共通部分系列

188

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める
  - 部分系列とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$の部分系列が $Z=\langle z_1,z_2,\ldots,z_k\rangle$であるとは，増加する添え字の列$\langle i_1,i_2,\ldots i_k\rangle$が存在し，$x_{i_j}=z_j$ を満たす．
  - 共通部分系列とは
    2つの系列の共通する部分系列

$X=\langle A, B, C, B, D, A, B\rangle$

$Y=\langle B, D, C, A, B, A\rangle$

$\langle B, C, A\rangle$ は共通部分系列

189

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める
  - 部分系列とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$の部分系列が $Z=\langle z_1,z_2,\ldots,z_k\rangle$であるとは，増加する添え字の列$\langle i_1,i_2,\ldots i_k\rangle$が存在し，$x_{i_j}=z_j$ を満たす．
  - 共通部分系列とは
    2つの系列の共通する部分系列

$X=\langle A, B, C, B, D, A, B\rangle$

$Y=\langle B, D, C, A, B, A\rangle$

最長共通部分系列(LCS)は$\langle B, C, A, B\rangle$と$\langle B, D, A, B\rangle$

190

問題

- DNAの塩基がどれぐらい一致しているか知りたい
- 最長部分系列を求める
  - 部分系列とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$の部分系列が $Z=\langle z_1,z_2,\ldots,z_k\rangle$であるとは，増加する添え字の列$\langle i_1,i_2,\ldots i_k\rangle$が存在し，$x_{i_j}=z_j$ を満たす．
  - 共通部分系列とは
    2つの系列の共通する部分系列
  - 最長共通部分系列問題とは
    $X=\langle x_1,x_2,\ldots,x_m\rangle$ と $Y=\langle y_1,y_2,\ldots,y_n\rangle$の
    最長共通部分系列(LCS)を求める

191

総当り作戦の破綻

- Xの部分系列の個数
  $\rightarrow O(2^m)$

総当りには$\Omega(2^m)$必要

192

アルゴリズムのステップ1: 最適解の構造を特徴づける

193

LCSの部分構造最適性

- $X=\langle x_1,x_2,\ldots,x_m\rangle$と$Y=\langle y_1,y_2,\ldots,y_n\rangle$のLCSが$Z=\langle z_1,z_2,\ldots,z_k\rangle$であると仮定すると

1. $x_m = y_n$ならば$z_k=x_m=y_n$であり，Z_{k-1}はX_{m-1}とY_{n-1}のLCSである
2. $x_m \neq y_n$のとき，$z_k \neq x_m$ならばZはX_{m-1}とYのLCSである
3. $x_m \neq y_n$のとき，$z_k \neq y_n$ならばZはXとY_{n-1}のLCSである

接頭語 $X_i=\langle x_1,x_2,\ldots,x_i\rangle$とする

194

LCSを求める手順

- $x_m = y_n$のとき
  - X_{m-1}とY_{n-1}のLCSを求める
- $x_m \neq y_n$のとき
  - 「X_{m-1}とYのLCS」と「XとY_{n-1}のLCS」の長いほうを採用する

部分問題の部分問題を共有している
→部分問題重複性

195

最適解の値

$c[i, j]$ = X_iとY_jのLCSの長さ

$$c[i,j]=\begin{cases}0 & (i=0\text{または}j=0\text{のとき})\\ c[i-1,j-1]+1 & (i,j>0\text{かつ}x_i=y_j\text{のとき})\\ \max(c[i,j-1],c[i-1,j]) & (i,j>0\text{かつ}x_i\neq y_j\text{のとき})\end{cases}$$

※問題に依存する条件により，検討すべき部分問題が変化する

196

アルゴリズムのステップ3: （ボトムアップに）最適解の値を求める

197

動的計画法の利用

- 再帰（メモ化なし）で計算すると指数時間
- 部分問題の個数は$\Theta(nm)$個のみ

→ 動的計画法で
ボトムアップ(昇順)に解く

LCS-LENGTH(X,Y)

```
m ← length[X]
n ← length[Y]
for i ← 1 to m
    do c[i,0] ← 0
for j ← 1 to n
    do c[0, j] ← 0
for i ← 1 to m
    do for j ← 1 to n
        do if x_i = y_j
            then c[i, j] ← c[i−1, j−1]+1
                b[i, j] ← "左上"
            else if c[i−1, j] ≥ c[i, j−1]
                then c[i, j] ← c[i−1, j]
                    b[i, j] ← "↑"
                else  c[i, j] ← c[i, j−1]
                    b[i, j] ← "←"
```

i=0, j=0を設定

左上から1つずつ計算

例 (p33)

	j	0	1	2	3	4	5	6
i		yj	B	D	C	A	B	A
0	xi	0	0	0	0	0	0	0
1	A	0						
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=0, j=0 を初期化

例 (p33)

	j	0	1	2	3	4	5	6
i		yj	B	D	C	A	B	A
0	xi	0	0	0	0	0	0	0
1	A	0	↑0					
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=1, j=1 のとき
$x_i \neq y_j$, $c[i\text{-}1,j] \geqq c[i,j\text{-}1]$

例 (p33)

	j	0	1	2	3	4	5	6
i		yj	B	D	C	A	B	A
0	xi	0	0	0	0	0	0	0
1	A	0	↑0	↑0				
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=1, j=2 のとき
$x_i \neq y_j$, $c[i\text{-}1,j] \geqq c[i,j\text{-}1]$

例 (p33)

	j	0	1	2	3	4	5	6
i		yj	B	D	C	A	B	A
0	xi	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0			
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=1, j=3 のとき
$x_i \neq y_j$, $c[i\text{-}1,j] \geqq c[i,j\text{-}1]$

例 (p33)

	j	0	1	2	3	4	5	6
i		y_j	B	D	C	A	B	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1		
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=1, j=4 のとき
x_i=y_j

204

例 (p33)

	j	0	1	2	3	4	5	6
i		y_j	B	D	C	A	B	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1	←1	
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=1, j=5 のとき
$x_i \neq y_j$, $c[i\text{-}1,j] < c[i,j\text{-}1]$

205

例 (p33)

	j	0	1	2	3	4	5	6
i		y_j	B	D	C	A	B	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1	←1	↖1
2	B	0						
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

i=1, j=5 のとき
x_i=y_j

206

例 (p33)

	j	0	1	2	3	4	5	6
i		y_j	B	D	C	A	B	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1	←1	↖1
2	B	0	↖1	←1	←1	↑1	↖2	←2
3	C	0						
4	B	0						
5	D	0						
6	A	0						
7	B	0						

207

例 (p33)

	j	0	1	2	3	4	5	6
i		y_j	B	D	C	A	B	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1	←1	↖1
2	B	0	↖1	←1	←1	↑1	↖2	←2
3	C	0	↑1	↑1	↖2	←2	↑2	↑2
4	B	0	↖1	↑1	↑2	↑2	↖3	←3
5	D	0	↑1	↖2	↑2	↑2	↑3	↑3
6	A	0	↑1	↑2	↑2	↖3	↑3	↖4
7	B	0	↖1	↑2	↑2	↑3	↖4	↑4

208

アルゴリズムのステップ4：最適解を構成する

209

PRINT-LCS(b,X,i,j)

if $i = 0$ **or** $j = 0$

  then return

if $b[i, j] =$ "↑←"

  then PRINT − LCS($b, X, i-1, j-1$)

    print x_i

elseif $b[i, j] =$ "↑"

  then PRINT − LCS($b, X, i-1, j$)

else PRINT − LCS($b, X, i, j-1$)

例 (p33)

i	j	0	1	2	3	4	5	6
		y_j	B	D	C	A	B	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1	←1	↖1
2	B	0	↖1	←1	←1	↑1	↖2	←2
3	C	0	↑1	↑1	↖2	←2	↑2	↑2
4	B	0	↖1	↑1	↑2	↑2	↖3	←3
5	D	0	↑1	↖2	↑2	↑3	↑3	↑3
6	A	0	↑1	↑2	↑2	↑3	↑3	↖4
7	B	0	↖1	↑2	↑2	↑3	↖4	↑4

コードの改善

- $b[i,j]$ を利用しない
  - $c[i, j]$の矢印の方向は$O(1)$ で求められる
  - 最適解の構成に要する時間: $O(m+n)$
  - メモリ量の節約: $\Theta(mn)$
- $c[i, j]$のメモリ量を削減する
  - 直前の行$c[i\text{-}1, j]$のみを覚えておけばよい
  - 最適解の値(長さ)のみを求めればよい場合
  - メモリ量の節約: $\Theta(mn)$

i	j	0	1	2	3	4	5	6
		y_j	B	D	C	A	B	A
0	x_i							
1	A							
2	B							
3	C	0	↑1	↑1	↖2	←2	↑2	↑2
4	B	0	↖1	↑1				
5	D							
6	A							
7	B							

演習: 以下のLCSを求めよ

i	j	0	1	2	3	4	5	6
		y_j	B	C	D	A	C	A
0	x_i	0	0	0	0	0	0	0
1	A	0						
2	C	0						
3	B	0						
4	D	0						
5	D	0						
6	A	0						
7	C	0						

演習: 以下のLCSを求めよ

i	j	0	1	2	3	4	5	6
		y_j	B	C	D	A	C	A
0	x_i	0	0	0	0	0	0	0
1	A	0	↑0	↑0	↑0	↖1	←1	↖1
2	C	0	↑0	↖1	←1	↑1	↖2	←2
3	B	0	↖1	↑1	↑1	↑1	↑1	↑2
4	D	0	↑1	↑1	↖2	←2	←2	↑2
5	D	0	↑1	↑1	↖2	↑2	↑2	↑2
6	A	0	↑1	↑1	↑2	↖3	←3	↖3
7	C	0	↑1	↖2	↑2	↑3	↖4	←4

例題:

最適二分探索木
(簡単バージョン)

問題の定義

入力
- n個の重みw_i $(1 \le i \le n)$が与えられる

出力
- $\sum_i w_i$ ×depth(i)
  を最小とする二分木（最適二分探索木）．
  ただしdepth(i)はi番目の葉の深さ

問題の直感的なイメージ
- 重みが探索される頻度に対応，データを二分木に整理するが，探索コストを少なくしたい

216

部分構造最適性

- 区間 $[i, j]$に関して，要素$i, i+1, \cdots, j$ の最適二分探索木の左部分木が要素$i, i+1, \cdots, k$を，右部分木が要素$k+1, k+2, \cdots, j$を含むとする
- この時，左右の部分木は，対応する要素の最適二分探索木になっているはず
- $c_{i,j}$ を，最適解の値とすると，
  $c_{i,j} = c_{i,k} + c_{k+1,j} + \sum_{l=i}^{j} w_l$

217

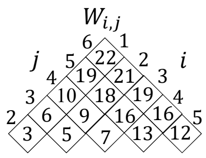
最適解を与える漸化式

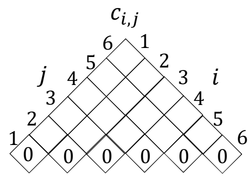
- $c_{i,j} = \min_{i \le k < j}\{c_{i,k} + c_{k+1,j}\} + W_{i,j}$
- ただし，$W_{i,j} = \sum_{l=i}^{j} w_l, c_{i,i} = 0$ とする．

218

演習

- 重みを 1, 2, 3, 4, 9, 3とする
- $c_{i,j}$ の表を埋めよ



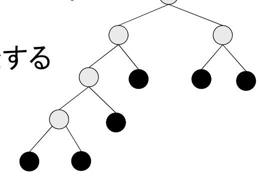


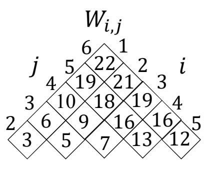
$c_{i,j} = \min_{i \le k < j}\{c_{i,k} + c_{k+1,j}\} + W_{i,j}$
ただし，$W_{i,j} = \sum_{l=i}^{j} w_l, c_{i,i} = 0$ とする．

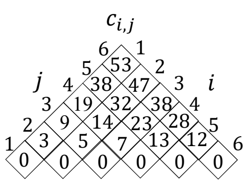
219

演習 (解答)

- 重みを 1, 2, 3, 4, 9, 3とする
- $c_{i,j}$ の表を埋めよ







220