

# Problem E

## お姫様の暗号解読

問題作成者: 野田  
解法作成者: 松本, 吉田  
問題文: 黄  
解説: 八森


# 問題概要

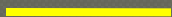
---

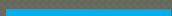
- ◎  $N$ 個の単語が ( $1 \leq N \leq 10$ ) 与えられる。  
 $N$  個の単語全部を部分文字列として含み、  
なおかつその長さが最小であるような文字列を求めよ。

# 例

◎ 単語:

abbd 

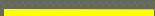
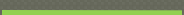

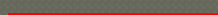
bab 

bdc 

dcab 

◎ 答え:

babbdcab

# Shortest Common Superstring

---

- この問題は、いわゆるShortest Common Superstring と呼ばれる。
- NP完全問題で、高速に解くためのアルゴリズムが無い。
- だが、 $N \leq 10$ なので、全探索しても間に合う。

# 全探索アルゴリズム(1/4)

---

- 前処理として、ある単語Aが別の単語Bの中に含まれていたら、Aを除去する。
- 前処理として、単語の末尾部と、別の単語の先頭部の共有文字数を計算する。（全ての単語の組み合わせに対して、この計算をし、 $N \times N$ 行列に保持。）

# 全探索アルゴリズム(2/4)

---

例:

文字列C: abbcd    文字列D: cdecff

abbcdecff

黄色い部分を共有している。

# 全探索アルゴリズム(3/4)

- 単語の全順列を生成して、各単語の順列に対して、単語を順番に連結していく。単語を連結するときには、片方の単語の共有部分を取り除く。
- 連結してできた文字列は、全ての単語を部分文字列として含んでいる。その中で、最短、辞書順で最小のものを保存し、出力。

# 全探索アルゴリズム(4/4)

---

◎ 計算時間:

$$O(N! * L)$$

$$= O(10! * 100)$$

$$\doteq O(300,000,000)$$

計算に時間はかかるけど、間に合います。

N: 単語の数

L: 単語の文字の総数



# DPで解くアルゴリズム(1/4)

- DPを用いれば高速に解ける。
- 前処理として、ある単語Aが別の単語Bの中に含まれていたら、Aを除去する。
- 前処理として、単語の末尾部と、別の単語の先頭部の共有文字数を計算する。（全ての単語の組み合わせに対して、この計算をし、 $N \times N$ 行列に保持。）
- ここまでは、全探索アルゴリズムと一緒に。

# DPで解くアルゴリズム(2/4)

- `table[N][2^N]`を確保。
- (最後に連結した単語のid, これまでに連結した単語の集合) に対し、連結してできる文字列で最短で辞書順最小な文字列を保持。
- 上の値から、  
(次に足す単語のid、  
これまでに連結した単語の集合  $\cup$  次に足す単語のid)  
が計算できる。
- 答えとなる値は、  
(最後に追加した単語のid(1~N)、全ての単語を含む集合)  
の[1~N]のどれかに入っている。

# DPで解くアルゴリズム(3/4)

- ◎ 単語同士の連結は~~右側から左側へ~~行う。  
左側から右側へ
- ◎ 文字列同士の辞書順比較を行いたい場合、  
最後に連結した~~単語同士の比較~~をすると  
よい。 文字列
- ◎ 前処理で、別の単語を含む単語を取り除いているので、最後に連結した文字列同士の比較で文字列の比較が正しく行える。

# DPで解くアルゴリズム(4/4)

- ◎ 計算時間: (N=10のとき)
- ◎  $O(2^N * N * L)$   
=  $O(2^{10} * 10 * 10)$   
≐  $O(1,000,000)$
- ◎ N ≤ 20までなら間に合う。

N: 単語の数

L: 一つの単語の長さ